Development of a method for optimal detection of emerging disease incursions

A thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy at Massey University

Shiyong Wang

2016

Submitted 29 January, 2016

Institute of Veterinary, Animal and Biomedical Sciences

Massey University

Palmerston North, New Zealand

Abstract

Emerging and re-emerging infectious diseases (ERID) are capable of generating sizable economic loss, and causing loss of life and social instability. To prevent and mitigate the negative impacts of ERID, it is imperative to have a sensitive surveillance system for early disease detection. Furthermore, from the economic perspective, resources are always scarce and have opportunity cost, so investment in surveillance programs has to demonstrate that it can maximize the utility of available resources. The thesis was focused on development and application of a software toolbox, **H**uman and **An**imal **Di**sease **Response** Program (HandiResponse), designed for (i) visualizing the disease risk landscape and representing spatial variation in the expected occurrence of a zoonotic disease both quantitatively and visually; (ii) evaluating economic benefit and costs of a single surveillance activity or a multi-component portfolio; (iii) identifying optimal use of resources for surveillance. It comprises four modules: (i) risk map development – HandiMap; (ii) surveillance portfolio development – HandiSurv; (iii) economic impact assessment – HandiEcon and (iv) surveillance optimization – OptiSurv.

The modules developed were tested on a number of data sets from various countries. The experience demonstrated that using satellite-derived data in combination with national statistical data to produce a disease risk map improved spatial prediction of avian influenza H5N1 outbreaks in southern Vietnam. Development of a risk map from satellite data for Crimean Congo Haemorrhagic Fever for Mongolia guided a field surveillance program which provided the first evidence that this disease is present in both animals and people in Mongolia. Finally an invented disease affecting pigs and people was used to investigate the likely consequences of an incursion of such a novel disease into Australia, involving both domestic and feral pigs and transferring to people. Risk-based and classical disease surveillance options were then tested for disease detection, and modelling work confirmed that a portfolio consisting of different options was the most technically and economically appropriate.

HandiResponse is a practical tool that could promote the implementation of risk-based surveillance approaches, and improve both technical and economic efficiency of surveillance programs for infectious diseases, particularly those affecting both people and animals.

Acknowledgements

Table of Contents

List of Figures

List of Tables

Abbreviations

Abbreviation	Explanation
ACF	Autocorrelation Function
AHP	Analytic Hierarchy Process
ARIMA	Autoregressive Integrated Moving Average
AUC	Area Under Curve
bTB	bovine Tuberculosis
CAC	Codex Alimentarius Commission
CCHF	Crimean Congo Haemorrhagic Fever
COS	Consequence of spread
CV	Coefficient of Variation
CWD	Chronic Wasting Disease
DALY	Disability Adjusted Life Year
DR	Direct Rating
EBL	Enzootic Bovine Leucosis
ELISA	Enzyme-linked Immunosorbent Assay
ERID	Emerging and Re-emerging Infectious Diseases
GARP	Genetic Algorithm for Rule-set Prediction
GIS	Geographic Information System
GLM	Generalized Linear Model
HandiEcon	Human and Animal Disease Economic Module
HandiMap	Human and Animal Disease Mapping Module
HandiResponse	Human and Animal Disease Response Program
HandiSpread	Human and Animal Disease Spread Program
HandiSurv	Human and Animal Disease Surveillance Module
HandiView	Human and Animal Disease View Program
HIV	Human Immunodeficiency Virus
HPAI	Highly Pathogenic Avian Influenza
IFA	Indirect Immunofluorescence Assay
IPPC	International Plant Protection Convention
LOS	Likelihood of Spread
MADM	Multi-Attribute Decision Making
MARP	Most at Risk Population
MCDA	Multiple-criteria Decision Analysis
NDVI	Normalized Difference Vegetation Index
OIE	World Organization of Animal Health
OptiSurv	Optimal Surveillance Module
PA	Point Allocation
PACF	Partial Autocorrelation Function
PCR	Polymerase Chain Reaction
POE	Probability of Exposure
PRRS	Porcine Reproductive and Respiratory Syndrome

1. Introduction

The research described in this thesis originated from my involvement in helping the affected countries and regions in implementing their responses to a number of major zoonotic diseases outbreaks and even epidemics. The first even was Severe Acute Respiratory Syndrome (SARS), for which I provided technical input when I worked at the Beijing Office of the World Bank. Subsequently I was responsible for managing the World Bank's Operations in China and Mongolia in response to the avian influenza H5N1 pandemic, and advised on the similar projects in India and Vietnam. I have also been involved in the research and control programs for HIV/AIDS early years in my career and then recently. These experiences led me to the view that it was necessary to have a mechanism or tools by which countries that were at risk of becoming infected (but which had limited expertise and resources) could design and implement appropriate surveillance and disease mitigation strategies. Such a perception was reinforced by my experience in spending much of the last year in Monrovia, to manage the World Bank's support to the national Ebola response in Liberia, and subsequently working on the design of a project to enhance disease surveillance capability in the West Africa Region.

The need for a comprehensive epidemiological and economic approach to both disease detection and subsequent control is imperative. However, the challenges for addressing both objectives at the same time, illustrated by Howe and his colleagues, are formidable (Howe, Häsler, & Stärk, 2013). It is beyond the scope of this PhD project to address all of them at once. Hence, in this thesis, I chose to focus only on the design and evaluation of a structured approach for assessing and identifying country-specific risk-based surveillance systems to detect an incursion of an emerging infectious disease in a given country.

The literature review in Chapter 2 demonstrates that there have been significant advances in the design of risk-based surveillance approaches. Unfortunately, there are few field experiences from developing countries. Such a dilemma calls for the development of a generic set of tools to facilitate the design of best-fit surveillance approaches for any emerging infectious diseases in any country.

The thesis hereafter documents the development of a novel toolbox called **H**uman and **An**imal **D**isease **Response** (HandiResponse) in order to bridge the gap between theory and practice of risk-

based surveillance approaches. The surveillance toolbox comprises the following four components: (i) HandiMap; (ii) HandiSurv; (iii) HandiEcon and (iv) OptiSurv, described in Chapters 3 and 4 of the thesis.

The first step in designing a risk-based surveillance approach that takes account of spatial variation in the probability of disease occurrence is to integrate and present the various risk factors proven or hypothesized to be responsible for or correlated with for disease occurrence. HandiMap was developed to produce an integrated risk landscape by combining a sub-set of global remote-sensed and national spatial data sets. In Chapter 5, three alternative risk maps were developed by using HandiMap, based on different epidemiological assumptions, to test the degree to which modelling H5N1 on these risk landscapes could improve the accuracy of spatial prediction of the disease outbreak in comparison with modelling on a "flat" risk landscape that took none of these factors into account.

In Chapter 6, I use a risk landscape approach to identify areas of Mongolia that are at either high or low risk of possible presence of Crimean Congo Haemorrhagic Fever (CCHF). Mongolia was considered free of the disease, although it occurs in neighbouring countries. As part of the disease investigation work undertaken by Mongolian scientists through an emerging diseases project that I managed, targeted cross-sectional surveys were undertaken, and serological evidence of CCHF virus circulation was found in both people and animals in epidemiologically determined high risk areas, but not in a low risk area. Thus the risk landscaping approach was of practical value in guiding surveillance activities, leading to the discovery that this disease was already present in Mongolia, but unrecognized.

The next step was to use risk data to develop an epidemiologically and economically appropriate combination of surveillance methods to detect a disease incursion in a country where the disease of concern would involve domestic animals, wild animals and people. No suitably comprehensive data set was available for an Asian country. Comprehensive spatial data was available on commercial and non-commercial pig herds in Australia and adequate evidence was available on movements within the industry. Spatial mapping of habitat suitability for feral pigs was also available for Australia, so in Chapter 7 this habitat map was used to establish a spatially defined population of feral pig families, which could interact with owned pigs in determining the spread of an invented zoonosis, called Austeria. The spread of the disease following a point incursion was simulated on this landscape, and eight alternative surveillance components were tested in a spatial modelling process for their ability to detect the incursion promptly and cost-effectively, each component having three levels of investigation intensity and three levels of detection sensitivity.

2

In Chapter 8, the findings from this modelling process are used in combination with economic data on the hypothesized effects of the disease on pig production and human health, to compare 100 million possible surveillance portfolios which could be put together from the 72 surveillance components and sub-components, and use the OptiSurv procedure that forms part of HandiResponse to identify the portfolios which best combine both prompt detection and cost-effective operation to detect the disease incursion.

In Chapter 9, the degree to which the project goals were achieved is assessed, and consideration is given to the next steps required to move from this "proof of concept" phase to a more comprehensive system which can be used by individual countries and the global community to detect and manage future emerging diseases.

1.1. References

Howe, K., B. Häsler, et al. (2013). "Economic principles for resource allocation decisions at national level to mitigate the effects of disease in farm animal populations." Epidemiology and infection 141(01): 91-101.

2. Review of Risk-based Surveillance for Infectious Diseases

2.1. Introduction

The emergence and spread of a series of major infectious diseases of zoonotic origin over recent years has led to a resurgence of interest in developing improved methods of responding to such problems, because they have major ramifications not just on health of people and animals, but also on economic growth, national development, social stability, and international movement of people and products (Calain, 2007; Gubler, 1998; Jones, Patel et al., 2008; McInnes & Lee, 2006; McMichael, 2004; States, 2000; UNDG, 2015). Infectious diseases have secondary effects on various aspects of social and economic life such as fertility, savings, investment, crop choices, food supply, human rights in terms of access to care and schooling, health care, migration decision, as well as animal welfare and product marketability (Piot, Muyembe, & Edmunds, 2014; Sachs & Malaney, 2002; Volkova, Bessell et al., 2011). The impact of infectious diseases is likely to grow in the future as a result of changing climatic conditions, with vector-borne diseases already widely distributed and causing serious health effects (Gubler, 2012) and economic effects (Gallup & Sachs, 2001), but expected to increase in geographical distribution, with newly important vector-borne diseases such as Zika virus adding to the challenges. The Ebola outbreak in West Africa during 2014 and 2015 caused economic stagnation and even recession in the three most severely affected countries. Other infectious diseases have also been expanding their geographic coverage (Anderson, Cunningham et al., 2004; Barrett, Kuzawa et al., 1998; Gubler, 2007; Kilpatrick, Chmura et al., 2006; Marano, Arguin, & Pappaioanou, 2007).

Surveillance is an essential public and animal health practice and one of the essential steps to counteract the increasing tide from infectious diseases (Abdullah, 2007; Bettcher, Sapirie, & Goon, 1997; Heymann & Rodier, 2001). Although there are varied definitions of surveillance for human and animal diseases, the common objectives of disease surveillance include:

- systematic monitoring of disease occurrence by population, place, and time;
- detection of unusual occurrence of disease or unusual epidemiological patterns;
- outbreak investigation and intensive follow-up to identify risk factors and potential points for

intervention; and

- implementing interventions and evaluating their effectiveness (Langmuir, 1963; Raska, 1966; Thacker & Berkelman, 1988; WHO, 2008).

In addition, surveillance has been used for documenting the health status of exported animals and animal products to demonstrate that they are free from diseases of trade concern (Kuiken, Leighton et al., 2005).

Despite its importance, implementation of infectious disease surveillance at national and global levels suffers from a number of political, technical and operational challenges resulting in divergence in methods and effectiveness of disease surveillance practices between developed and developing economies (McInnes & Lee, 2006). Some of the frequently cited deficiencies include: (i) inadequate collaboration between human and animal health personnel (Tsai, Scott et al., 2009); (ii) disincentives associated with disease outbreaks; cost and time needed for surveillance (Alban, Pozio et al., 2011; Alleweldt, Upton et al., 2009; FAO, OIE et al. 2008; Wagner, Moore, & Aryel, 2011); (iii) inadequate capability in terms of insufficient health workforce and lacking of adoption of a multi-disciplinary approach (Barbiero, 2014; Gubler, 2012; Witt, Richards et al., 2011); (iv) data issues such as inaccessibility or poor integration of data sources (Woolhouse, 2011) and (v) deficiencies in IT and laboratory infrastructure (Tsai, Scott et al., 2009). The points mentioned above highlight the need and urgency for development and utilization of accessible and accountable surveillance approaches which have both high technical efficiency and high economic benefit.

Risk-based surveillance (RBS) is a relatively recently adopted approach which can potentially contribute significantly to satisfying these requirements. It has been used for human and animal disease surveillance, food safety and environmental contamination monitoring (Chon, Ohandja, & Voulvoulis, 2012; Gkogka, Reij et al., 2011; Stark, Regula et al., 2006).

This review aims to review risk-based surveillance practices and their utility, identify enabling factors as well as the challenges for promoting RBS. It will also recommend potential enhancements that would make RBS more widely available and practical to apply in resource-poor countries, and outline research needed to achieve this objective. This will then lead into the remainder of the thesis, which describes research on this theme.

2.2. Methods

2.2.1. Literature search strategy

The search included the following topics:

- A.   infectious disease or communicable disease surveillance or public health surveillance or bio-surveillance or early warning or disease reporting;

B.	risk-based approach or optimization or cost effectiveness or prioritization or targeted or sentinel surveillance;

C.	social network analysis.

Searching queries were "A AND B; "A AND C". This search included an iterative process to refine the search strategy by testing several search terms and incorporating new search terms when new relevant citations were identified. The databases searched included Web of Science, MEDLINE, Current Contents Connect, Biological Abstracts, Centre for Agriculture and Bioscience International (CABI) database and Food Science and Technology Abstracts (FSTA).

Besides, additional relevant articles have been added on the list for review after citation cross checks by using Google scholar searching engine (up to 15 pages).

2.2.2. Inclusion criteria

Refinements of the search strategy included focusing on infectious disease, tropical medicine, social sciences, mathematics, medical informatics, veterinary sciences, zoology, and entomology. Only publications in English published since 1980 and with access to full document were included in the review. The hard criterion used for final inclusion is containing descriptions on risk-based approach for surveillance.

2.3. Results

2.3.1. General information on the review

65,738 hits met the initial inclusion criteria. By reviewing the topic, 491 publications were selected for further content review. Of them, 280 were included in the final review and analysis according to their content.

The early use of a risk-based approach for infectious disease surveillance could be dated back to early 1980s involving risk mapping and social network analysis to identify high risk groups for HIV surveillance (Klovdahl, 1985; Shannon, 1981). The other two early studies were about using a scenario tree approach for demonstration of freedom from disease and sentinel surveillance of HIV-1 among pregnant women (Kigadye, Klokke et al., 1993; Sergeant, Cameron et al. 1990). The number of publications relevant to RBS have been increasing over the years, particularly since 2005 (Figure 2-1). A majority of publications on RBS included in this review were from Web of Science and Google Scholar (Figure 2-2). By origin of publications, Europe contributed 32 percent, following by North America (20%). Only 15 percent of reviewed publications were from developing countries (Figure 2-3). Out of 280 reviewed, seventy percent were on risk-based sampling, thirteen percent on risk-based requirement and twelve percent on risk-based prioritization.

7
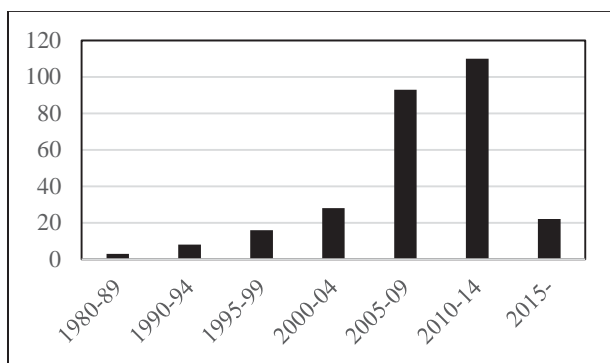
Figure 2-1. Publications on RBS by year

Figure 2-2. Publications on RBS by source

Note: CABI, Current Contents Connect, Biological Abstracts, Centre for Agriculture and Bioscience International; FSTA, Food Science and Technology Abstracts; GS, Google Scholar; WOS, Web of Science.
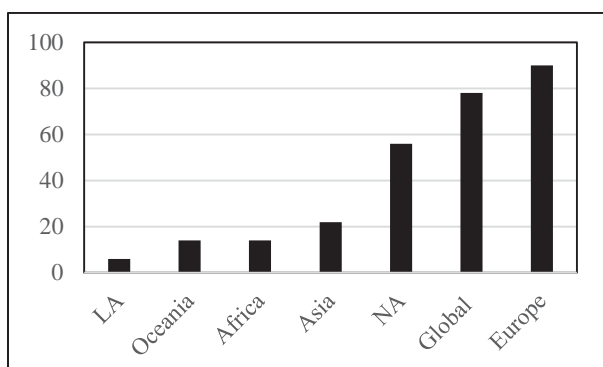
Figure 2-3. Publications on RBS by continent
Note: LA, Latin America; NA, North Africa.
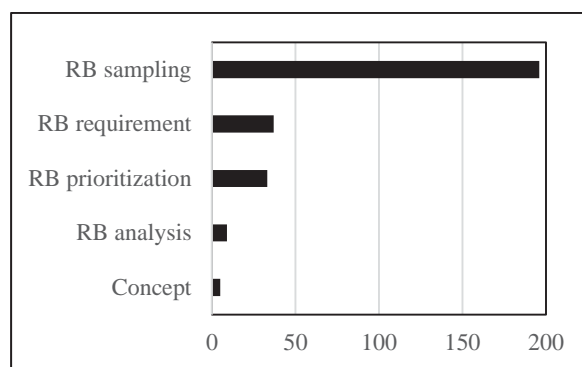
Figure 2-4. Publications on RBS by approach
Note: RBS, risk-based surveillance. X-axis represents the number of the peer reviewed articles.

258 publications clearly indicated on what health issues they were focused. Thirty three percent were on human infectious diseases, 32% on animal infectious diseases and 26% on zoonoses. Other health topics covered include food safety, wildlife disease, anti-microbial resistance, environmental pollution as well as vectors for infectious diseases (Figure 2-5).
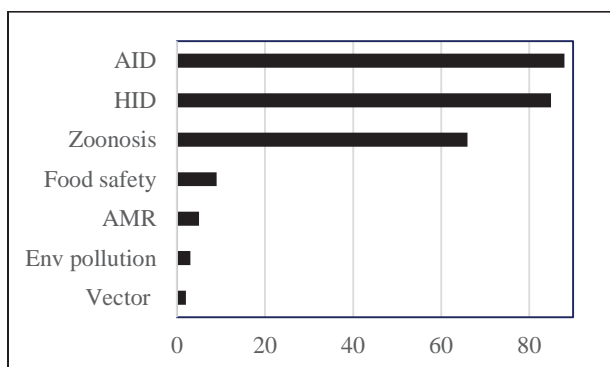
Figure 2-5. Focus of Risk-based Surveillance
Note: AID, animal infectious disease; AMR, anti-microbial resistence; Env, environment; HID, human infectious disease. X-axis represents the number of the peer reviewed articles.

2.3.2. Definition and objectives of risk-based surveillance

For risk-based approach, risk is the probability of occurrence and severity of consequences caused by a hazard (Stark, Regula et al. 2006). Some key definitions for RBS include:

1. A surveillance program in the design of which exposure and risk assessment methods have been applied together with traditional design approaches in order to assure appropriate and cost-effective data collection (Stark, Regula et al., 2006).

2. Use of information about the probability of occurrence and the magnitude of the biological and/or economic consequence of health hazards to plan, design and/or interpret the results obtained from surveillance (Hoinville, Ronello, & Alban, 2011).

3. The application of qualitative and quantitative methods to increase surveillance efficiency by directing surveillance activity to:

- the population of interest based on exposure to factors that may predispose it to disease or infection, or

- subpopulations where due to host factors, the disease or infection is most likely to be found, or

- prioritizing populations where the consequences of disease or infection could be severe.

A common thread shared by these definitions is the emphasis on prioritization by using risk-based approach in order to achieve either higher technical or economic efficiency of surveillance program. The rationale underpinning the risk-based strategies is that higher risks merit higher priority for surveillance resources as such investments would yield higher benefit (Stark, Regula et al. 2006).

2.3.3. Typology of RBS

Considerable heterogeneity exists in the approach and methodology for risk-based surveillance (Reist, Jemmi, & Staerk, 2012). At least three types of RBS were summarized by Stark et al. They are hazard selection, selection of population strata and sample size calculation (Stark et al., 2006). Hoinville et al. have further elaborated RBS practices into four groups: risk-based prioritization, risk-based requirement, risk-based sampling and risk-based analysis (Hoinville, Alban et al. 2013).

There are two prominent RBS variants: the first one is targeted surveillance, which is defined more generally as surveillance focusing on sampling high-risk populations. Some scholars recommended to reclassify it as risk-based sampling (Hoinville, Alban et al. 2013). The second one is sentinel surveillance, that focuses on specific high risk subpopulations, or animals instead of any human populations, to obtain timely information in a relatively inexpensive manner rather than to derive precise estimates of prevalence or incidence in the general population (McCluskey, 2003b).

This review results were presented by following the categorization method proposed by Hoinville, et al.

2.3.3.1. Risk-based prioritization

Risk-based prioritization is defined as an approach for determining which hazards should be selected for surveillance, based on information about the probability and the extent of (biologic and/or economic) consequences of their occurrence (Hoinville, Alban et al. 2013).

There are about 1,415 known species of pathogens that cause human infectious diseases (Pedley & Pond, 2003; L. H. Taylor, Latham, & Mark, 2001). Their occurrence varies in magnitude, severity and change with time. Prioritization among them is motivated by the need to ensure that scarce resources could be used on the most important ones to attain the highest benefit in improving health and welfare of human and animals.

Multi-attribute decision making (MADM) is one of the most commonly used disease prioritization methods. This structured approach ranks or groups alternatives from a finite set of discrete decision alternatives based on comparisons by using predefined criteria. In the case of disease prioritization, disease prevalence and disease incidence, case fatality, disease adjusted life year (DALY) and preventability, etc. are some of the criteria that have been applied. Criterion weights are described separately by evaluating the trade-offs that decision-makers are prepared to make between them (Brookes, Vilas, & Ward, 2015). MADM generally involves the following four distinctive steps:

1. structuring the decision problem;
2. assessing possible impact of diseases;
3. determining preferences (values) of decision−makers; and
4. evaluating and comparing disease priorities.

The outcome is that diseases will be ranked according to the weighted score of the summed products of a subjective weight and an objective/subject measurement by each criterion for each disease (Brookes, Hernández-Jover et al. 2015).

MADM usually is conducted by enlisting a group of topic experts/key informants and using Delphi method to elicit their judgements (Cediel, Villamil et al., 2013). Analytic Hierarchy Process (AHP) has been used to help weight elicitation (Kadohira, Hill et al., 2015; Ricci, Capello et al., 2013). The consistency between the two rounds of assessments by the experts could be checked by Spearman rank correlation coefficient (Brookes, Hernández-Jover et al. 2015) and validity of the approach can be tested through sensitivity analysis (Ricci et al., 2013). Two highlights emphasized by the users of such an approach are transparency and reproducibility (Rushdy & O'Mahony, 1998).

For human disease, a typical example of MADM is the weighted score system for prioritizing infectious diseases for surveillance purpose developed by Krause (Krause, 2008b). The system included twelve criteria under the four categories: (a) burden of disease (b) epidemiological dynamic;

(c) information need and (d) health gain opportunity. Such a method was developed based upon reviewing earlier works (Doherty, 2000; WHO, 2003; Weinberg, Grimaud, & Newton, 1999). Score for a disease is calculated by using linear weighted sum. The detailed process of such an approach is illustrated in Figure 2-6.
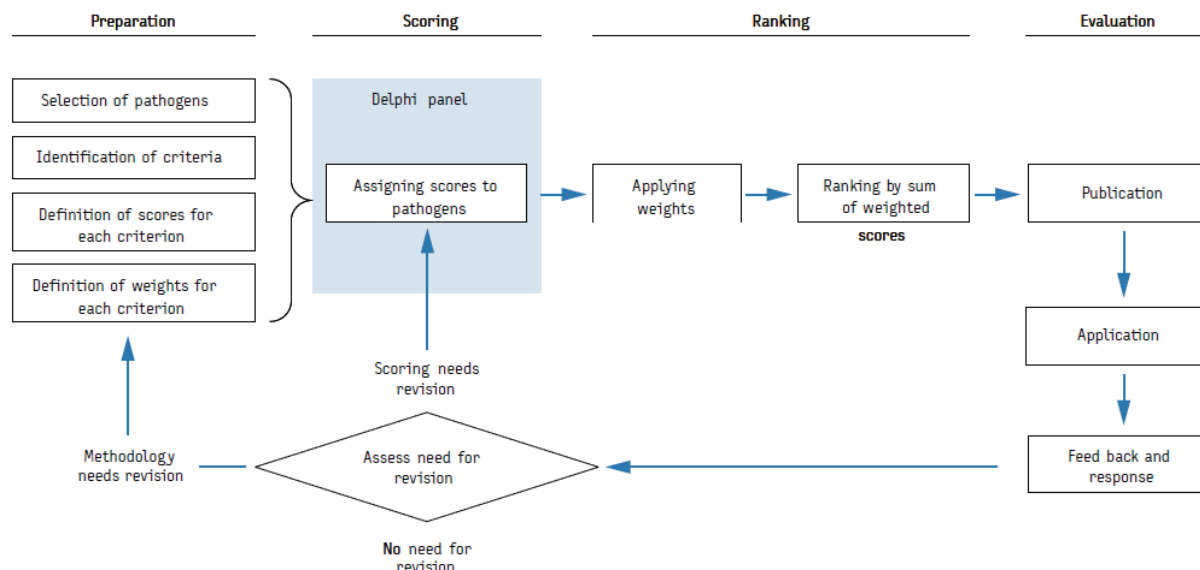
Figure 2-6. Algorithm for Prioritization of Disease Surveillance
Note: Krause, G. (2008). Prioritization of infectious diseases in public health-call for comments.

For animal infectious disease, risk estimation is based on the probability of occurrence and severity of consequences. The final risk score can be quantified by using the following formulae:

$$Final\ score = \Sigma weighted\ POE \times \Sigma weighted\ LOS \times \Sigma weighted\ COS$$

Where POE stands for probability of exposure, LOS, likelihood of spread and COS, consequence of spread. Standardization is sometimes used to make sure the scores are comparable when numbers of criteria are different (McKenzie, Simpson, & Langstaff, 2007; Stebler, Schuepbach-Regula et al., 2015).

MADM has been used for priority setting for surveillance of human and animal infectious diseases (Balabanova, Gilsdorf et al., 2011; Cardoen, Van Huffel et al., 2009; Ciliberti, Gavier-Widen et al.,2015; Cox, Revie, & Sanchez, 2012; Doherty, 2000; East, Wicks et al., 2013; Gilsdorf & Krause, 2011; Gustafson, Klotins et al., 2010; McKenzie, Simpson & Langstaff et al., 2007; Paige, Chaudry, & Pell, 1999; Robinson, Burgman, & Cannon, 2011; Weinberg, Grimaud & Newton, 1999), food safety (Presi, Stärk et al., 2009; Ricci, Capello et al., 2013; Tavernier, Dewulf et al., 2011), environmental pollution (Chon, Ohandja & Voulvoulis, 2012), as well as disease prioritization related to climate change (Akin, Martens, & Huynen, 2015; Cox, Revie & Sanchez, 2012), international trade

11

(Brookes, Hernández-Jover et al., 2014), tourism (Economopoulou, Kinross et al., 2014) as well as concerns over bioterrorism (Ryan, 2008).

One limitation related to MADM is how to demonstrate the validity of such an approach. Such an issue arises because of a number of reasons. Firstly, there is uncertainty and variability associated with disease impacts. One solution is to conduct sensitivity analysis to assess the robustness of prioritization results due to variability in inputs and heterogeneity in preferences for the importance of criteria among evaluators, decision makers, or stakeholders (Cediel, Villamil et al., 2013). This can be dealt with by using objective metrics to represent disease impacts such as incidence, disability adjusted life years, etc. Probabilistic inversion and conjoint analysis were suggested to be used for deriving criterion weights. Conjoint analysis quantifies the variation among decision making participants to allow generalization of the results to the wider population (Ng & Sargeant, 2013). Probabilistic inversion is a statistical method to infer weights for criteria from a large number of participants (Brookes, Hernández-Jover et al., 2014; Neslo & Cooke, 2011). A disadvantage of probabilistic inversion and conjoint analysis is that they rely on statistical methods to produce valid results, hence potentially require large numbers of participants (Brookes et al., 2015). Thirdly, diseases are constantly changing and so is the decision makers' preference on criteria. Hence, it is recommended that disease prioritization be updated periodically (Brookes, Vilas & Ward, 2015).

2.3.3.2. Risk-based sampling

Risk-based sampling is defined as designing a sampling strategy to reduce the cost or enhance the accuracy of surveillance by preferentially sampling strata (e.g. age groups or geographical areas) within the target population that are more likely to be exposed, affected, detected, become affected, transmit infection, or cause other consequences (Hoinville, Alban et al., 2013).

Information on relative risk, which is key for RBS approaches, can be generated through conventional epidemiological investigations such as case control studies, and cohort studies (Calvo-Artavia, Nielsen et al., 2013; Kung, Morris et al., 2007; Winkelstein, Lyman et al., 1987). For this review, the attention was focused on some novel approaches for generating information on relative importance for different strata of disease under study that are described below.

Social network analysis (SNA)

Infectious disease transmission involves direct interaction(s) between an infected host and a susceptible individual except in certain situations that infectious diseases can be transmitted by certain vector (e.g. mosquitos) or via airborne, or contaminated environment (Klovdahl, 1985). Contacts between individuals, either homogeneous or heterogeneous, can be represented by different network topologies. Simulation studies demonstrate that epidemic size and mean time to maximum size vary between different network topologies (Christley, Pinchbeck et al., 2005; Fevre, Bronsvoort et al., 2006; Levin, Grenfell et al., 1997; Shirley & Rushton, 2005). Analyzing geographic connections or

functional interactions among a social network can aid in understanding the nature and spread of a disease so as to inform decision making on disease management (Klovdahl, 1985; Smieszek, Fiebig, & Scholz, 2009).

Carroll et al. defined social network as a graphical representation of social relations or exposures consisting of nodes (individuals within the network) and ties/link (relationships between individuals) (Carroll, Au et al. 2014). SNA is a strategy for investigating social structures through the use of network and graph theories (Otte & Rousseau, 2002; Wasserman & Faust, 1994). Analysis based upon graphic theory has gained increasing popularity (Fiebig, 2011).

The major steps of a comprehensive SNA practice comprise: data collection and preparation, building a network and analysis and testing hypotheses. This is shown in the illustration published by Farine, et al (Figure 2-7) (Farine & Whitehead, 2015).

Figure 2-7. Primary steps and key considerations in the collection and analysis of animal social networks

For visualization or sociogram development, social network analysis uses shape, colour, direction and position to convey information on features of individuals and groups, clusters or paths. About 30 free or commercial software programs have been developed to facilitate and standardize social network analysis. By comparing functionality, support and user friendliness, UCInet and NetMiner have been highly recommended for social network analysis (Huisman & Van Duijn, 2005)

Metrics commonly used for social network analysis include:

- relative degree, closeness and betweenness centrality of nodes (Ortiz-Pelaez, Pfeiffer et al., 2006);
- components such as strong component or weak component (Grange, Van Andel et al., 2014; Newman, 2003; Rautureau, Dufour, & Durand, 2011; Robinson & Christley, 2007);
- ingoing and outgoing infection chains (Noremark, Hakansson et al., 2011) and
- clustering coefficient (Grange, Van Andel et al., 2014; Ribeiro-Lima, Enns et al., 2015).

SNA typically represents interactions on static networks. However, the social network could be subdivided to represent contacts or movements occurring during different time periods. Using temporal variation including seasonality of centrality measures were practiced to inform disease surveillance (Hamede, Bashford et al., 2009; Noremark, Hakansson et al., 2011; Sanchez-Matamoros, Martinez-Lopez et al., 2013).

A majority of SNA related publications in the human health sector have been on identification of high risk contacts (Christley, Pinchbeck et al., 2005; Eames, Tilston et al., 2012; Klovdahl, 1985; Morris, Zavisca, & Dean, 1995) and associated behaviours (Eames & Keeling, 2002), etc. Evidence on the differences of these measured parameters by subpopulation has been used to improve disease surveillance activities. Use of SNA in preventive veterinary medicine has been increasing over the years (Brooks-Pollock, Roberts, & Keeling, 2014; Cumming, Hockey et al., 2008; Dubé, Ribble et al., 2009; Fevre, Bronsvoort et al., 2006; Ortiz-Pelaez, Pfeiffer et al., 2006; Ribeiro-Lima, Enns et al., 2015; Wiratsudakul, Paul et al., 2014).

Using SNA for poultry movements, Martin, et al identified that the values of mean degree and k-neighbours of nodes (poultry farms) in Highly Pathogenic Avian Influenza (HPAI) H5N1 infected counties were significantly higher than those in non-infected counties in southern China. Besides, HPAI infected live bird markets had higher mean degree and k-neighbours of nodes with other infected markets compared with the same parameters for uninfected markets (Martin, Zhou, et al., 2011). The study confirmed the epidemiological importance of live bird markets in disease transmission such that they should be targeted for HPAI surveillance. Ribeiro-Lima, et al. analyzed four years' of cattle movements in the bovine tuberculosis (bTB) accredited areas in Minnesota, USA. Based upon the data from SNA, they calculated a risk score for each farm and categorized them into high, medium and low risk groups. The study discovered that the higher risk group, although only 14 percent of the all farms,

corresponded to 80 percent of the cumulative risk for the farms in the bTB area (Ribeiro-Lima, Enns et al., 2015). Use of SNA to identify high risk movements was also documented in other studies (Cumming, Hockey et al., 2008; Fevre, Bronsvoort et al., 2006; Fiebig, 2011; Martinez-Lopez, Perez, & Sanchez-Vizcaino, 2009; Ortiz-Pelaez, Pfeiffer et al., 2006; Sanchez-Matamoros, Martinez-Lopez et al., 2013; Van Kerkhove, Vong et al., 2009).

Some critical assumptions for a credible SNA identified by the review are:

- SNA assumes that the network underlying the study is a complete one in that all the possible contacts and relations between members in the network are included, hence transmission occurs only due to existing links in the network (Ortiz-Pelaez & Pfeiffer, 2008). However, a closed population (complete network) is a rare situation (Christley, Pinchbeck et al., 2005);
- SNA assumes the observed network represents the real world network. This may also not be true. Such an issue will be further discussed later. However, the more the credible data available on the network under study, the better the observed network will resemble the real one.

A quite small number of publications have focused on how to improve the quality of SNA analysis. For instance, checking the validity of the SNA results, using bootstrapping or jackknifing was suggested to estimate the confidence intervals around network measurements (Lusseau, Whitehead, & Gero, 2008; Whitehead, 2008). To make sure the observed network mirrors the real network (defined as a correlation between the edges of the real and the observed network of at least 0.8), Whitehead provided a guideline for estimating the sample size. For instance, a network that is moderately socially differentiated, where the coefficient of variation (CV) of edge weights of the real network is approximately 0.2, requires a mean of about 50 identifications per pair of connected nodes. This decreases as the network becomes more strongly differentiated, for example as relationships become less mixed and start to resemble pairs forming territories (Whitehead, 2008). Besides, based upon a simulation study, Franks, Ruxton and James suggested that increasing the frequency of studies, rather than increasing the proportion of individuals sampled in each study generates a more robust network when the social network under investigation is stable (Franks, Ruxton, & James, 2010).

Data quality and availability could present a challenge for SNA in the ways are illustrated below:

- Complete and reliable data on disease relevant contacts and movements might be unavailable, in particular in the early phases of a disease outbreak. This points out such pieces of information need to be collected systematically in advance of a disease outbreak and updated regularly. Development of templates to collect field intelligence in a structured way during outbreak investigations could also be an option (Ortiz-Pelaez & Pfeiffer, 2008).
- New sources and modalities for social network information collection need to be tapped. For instance, technological advances in human (mobile phone tracking, wearable sensors) and animal

tracking (proximity loggers, FRID tags) are rapidly increasing the amount of data collected (Eagle, Pentland, & Lazer, 2009; Isella, Romano et al., 2011; J. Krause, Krause et al., 2013).

- The information on social networks usually are collected over a short period of time and represent snapshots of interactions between network members, while the dynamics of any given networks is constantly changing. How to best use such information to understand the disease spread over periods of months or years is still an unanswered question (Bansal, Read et al., 2010);

- Data on networks are self-reported in most cases, which might suffer from recall bias and missing information (Smieszek, Burri et al., 2012). To overcome these issues, internet based surveys, automated data collection, electronic self-administered questionnaires were suggested and practiced (Eames, Tilston et al., 2012; Prah, Copas et al., 2013; Salathé, Kazandjieva et al., 2010).

- Last but not least, the large size of some databases and the subsequent high volume of data could overwhelm currently available SNA software. One practical suggestion is to aggregate detailed data into a larger epidemiological unit that serves a node (Ortiz-Pelaez, Pfeiffer et al., 2006).

Disease mapping

Spatial epidemiology focuses on at least four issues: (i) disease mapping, (ii) disease clustering; (iii) geographical correlation analysis and (iv) testing hypothesis (Berke, 2004; Carroll, Au et al., 2014). Disease mapping is often an exploratory analysis used to get an impression of the spatial distribution of a disease or its corresponding risk factors (Abrial, Calavas et al., 2003; Berke, 2004; Craig, Sharp et al., 2007; Cringoli, Rinaldi et al., 2005; Lawson, Biggeri et al., 1999; Mak, Morshed, & Henry, 2010; Noor, Kinyoki et al., 2014; Samat & Ma'arof, 2014; Shannon, 1981).

With the development of openly accessible remote sensing (RS) data, geographic information system (GIS) techniques and spatial statistics, risk surfaces containing information on the relative importance and the clustering of disease occurrence, at-risk populations, disease risk factors, vectors, reservoirs and environmental determinants can be projected as continuous maps either dynamically or statically (Abrial, Calavas et al., 2003; Fuller, Trevon et al., 2011; Grenfell, Bjørnstad, & Kappey, 2001; Hay, Guerra et al., 2009; Keeling, Woolhouse et al., 2001; Kitron, 2000). Disease maps can be produced in different granularities so that global, regional, country and even village level surveillance activities can be informed (Craig, Sharp et al., 2007; Gemperli, Sogoba et al., 2006; Hay, George et al., 2013; Noor, Kinyoki et al., 2014; Ribeiro, Seulu, et al., 1996).

Risk maps have been used to inform or refine RBS approaches; for example, monthly risk maps for bluetongue (BT) were developed in Switzerland by combining two components: (i) the monthly *Culicoides* vector habitat suitability maps which were based on temperature, humidity and altitude environmental descriptors, and (ii) monthly R_0 (basic production rate) maps (Racloz, Venter et al., 2008). The risk maps identified high risk locations and high risk months during a year and the

information generated from these maps was used for guiding the national targeted sentinel BT surveillance in the country (Racloz, Venter et al., 2008).

Martin et al. demonstrated the cyclic process between risk mapping and risk-based surveillance in a study on HPAI H5N1 in China. Reported poultry HPAI H5N1 outbreaks or evidence on existence of HPAI H5N1 virus at bird markets were used as dependent variables. Seven explanatory variables were selected based upon literature review. They used bootstrapped logistic regression and boosted regression trees (BRT) to estimate the relationship between disease/infection status and the selected risk factors and environmental descriptors. The risk maps turned out to have high prediction power. The study revealed that distribution of HPAI H5N1 risk in China appeared more limited geographically than previously assessed. This information could be used for a better targeted surveillance program (Martin, Pfeiffer, et al., 2011). A similar approach of combining geographic information system (GIS) techniques and statistical methods was used for understanding the relationship between environmental descriptors such as temperature, Normalized Difference Vegetation Index (NDVI) and precipitation with human Monkey Pox virus disease. The study discovered that proximity to dense forest and the habitat preferred by rope squirrels were two key risk factors for the occurrence of human Monkey Pox cases in Sankuru district in Democratic Republic of Congo. The risk of contracting the disease was significantly greater near sites predicted to be suitable habitat for squirrels (OR = 1.32; 95% CI 1.08–1.63). Semi-deciduous rainforests with oil-palm, the rope squirrel's main food source, was recommended as the basis for prioritizing surveillance for Monkey Pox disease (Fuller, Trevon et al., 2011).

Hay et al. summarized a disease mapping process using GIS and statistical methods in the following steps:

1. to define the definitive extent of the disease with the data on disease occurrence collected from various sources such as literature, internet, Gen Bank, etc.;
2. to infer pseudo-absence points with the definitive extent and occurrence point data;
3. to use statistical techniques to characterize points of presence and pseudo-absence against the range of explanatory variables; and
4. to predict the probability that the disease occurs at each location and thereby generate a risk map with a quantified measure of uncertainty by using the relationships between points of presence and pseudo-absence and explanatory variables (Figure 2-8) (Hay, George et al., 2013).
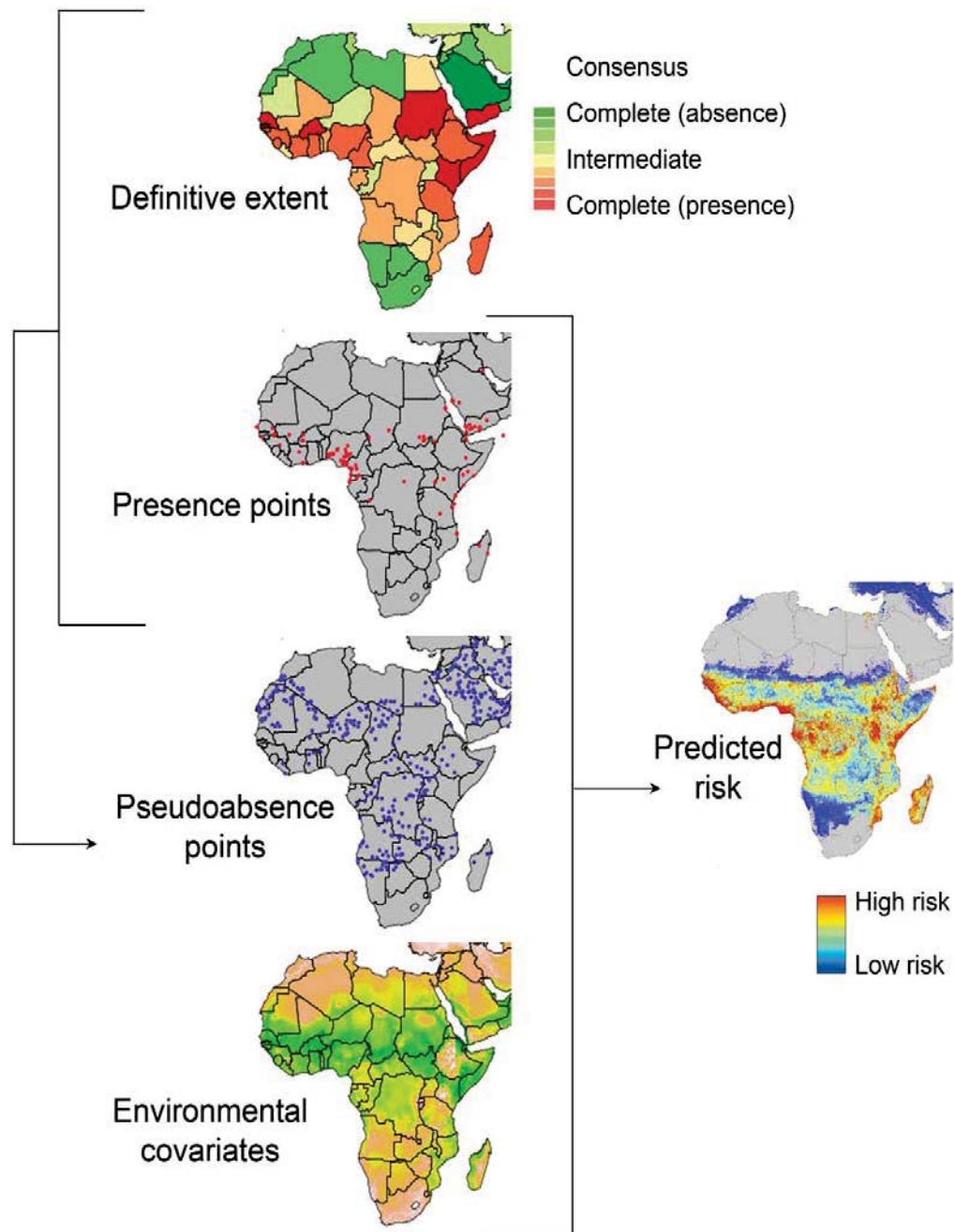
Figure 2-8. A schematic overview of the process of predicting spatial disease risk
Note: Hay, S. I., D. B. George, et al. (2013). "Big Data Opportunities for Global Infectious Disease Surveillance." Plos Medicine 10(4).

Risk mapping can provide early warning of an increase in disease cases or an outbreak. By using cubic spline function to assess a non-linear exposure and response association between weather predictors and dengue cases, Hii et al. estimated relative risks of dengue cases as the function of weekly mean temperature and cumulative rainfall. The study ascertained that increase in weekly mean temperature and cumulative rainfall precedes the increase in reported dengue cases by 4 to 20 and 8 to 20 weeks respectively (Hii, Rocklov et al., 2012). Similarly, another study using sea surface temperature and

NDVI as the predictor variables for early warning of Rift Valley Fever, the incidence of Rift Valley Fever outbreak could be forecast up to 5 months in advance (Linthicum, Anyamba et al., 1999).

Using environmental or climatological descriptors for disease risk mapping has gained popularity. This is particularly true for mapping vector borne diseases (Beck, Lobitz, & Wood, 2000; Kalluri, Gilruth et al., 2007; Rogers, 2006; Rogers, Randolph et al., 2002; Rogers, Tucker, & Myers, 2002). Most mapped vectors include mosquitos, ticks, black flies, tsetse flies and sandflies (Table 2-1). Further, availability of environmental satellite data for mapping infectious diseases was reviewed and summarized by Hay, et al. (Hay, Tatem et al., 2006).

Table 2-1. Key vectors, relevant environmental/climatological predictors and diseases

Vector	Environmental/climatological descriptors	Examples of VBDs
Mosquito	• NDVI, NDVI variability, leaf area index • Temperature • Elevation • Distance to waterways • Rainfall, sea surface temperature variations, cold cloud duration	• Rift valley fever • Dengue fever • Malaria • West Nile virus • Yellow fever
Tick	• NDVI, land cover • Temperature • Elevation • Rainfall	• Lyme disease • Crimean-Congo Hemorrhagic Fever
Black fly	• Land cover	• Onchocerciasis
Tsetse fly	• NDVI • Land surface temperature • Cold cloud duration; • Elevation	• Sleeping sickness
Sandfly	• NDVI	• Visceral Leishmaniasis
Rodent	• NDVI • Rainfall • Temperature • Elevation	• Lassa fever • Plague

Source: Kalluri, Satya, et al. "Surveillance of arthropod vector-borne infectious diseases using remote sensing techniques: a review." PLoS Pathog 3.10 (2007): 1361-1371; Fichet-Calvet, Elisabeth, and David John Rogers. "Risk maps of Lassa fever in West Africa." PLoS Negl Trop Dis 3.3 (2009): e388.

In a review, Eisen introduced four groups of spatial and space-time risk models for risk mapping, including:

- kriging for producing smooth interpolated maps for dependent variables;

- generalized linear models (GLM) to identify environmental or socioeconomic predictors for risk of exposure to disease and develop continuous spatial surfaces that present estimates of risk for exposure to vectors or pathogens. The results can be used to guide disease surveillance, prevention and control. Other models include generalized additive models and Bayesian approaches. Using these models, risk surface can be extrapolated to non-surveyed locations with ecological and climatic characteristics similar to those of the model development area;

- presence-only machine learning (rule-based) algorithms or dynamic simulation models. Examples include genetic algorithm for rule-set prediction (GARP) (Stockwell, 1999), MAXENT (Phillips, Anderson, & Schapire, 2006) and a machine learning algorithm based on maximum entropy and simulation models such as CLIMEX (Sutherst, Maywald, & Kriticos, 2007). Comparison of these models was discussed briefly. For instance, it was claimed that surface developed by using GARP may be useful for identifying new areas where diseases are likely to emerge, while surface produced by generalized linear model is better for guiding resource allocation to surveillance, prevention and control;

- space-time risk models. These models are useful to detect changing risk patterns hence their outputs may be suitable for early-warning systems. Commonly used methods include space-time permutation scan statistics (e.g., SaTScan), Knox tests, generalized additive mixed models and Bayesian hierarchical regression models. All these models highlight the importance of quality environmental, biological and epidemiological data, the most important prerequisite for a good modelling work (Eisen & Eisen, 2011).

Despite the fact that infectious disease risk maps would be valuable to policy makers prioritizing limited resources, out of 355 infectious diseases of clinical significance, only 4% of them have been mapped comprehensively (Hay, Battle et al., 2013). Obstacles responsible for the gaps frequently mentioned include:

- laborious steps in primary data acquisition, processing and positioning,
- the lack of collaboration between RS scientists and biologists;
- unavailability of sophisticated, statistical GIS;
- challenges in selection of best model(Robertson, Nelson et al., 2010);
- unavailability of georeferenced and spatially explicit disease data; (vi) inaccessibility to high resolution and low cost imagery;
- data inconsistencies in spatial, spectral, and temporal resolutions among satellite sensors as well as data heterogeneity (Gao, Mioc et al., 2008; Herbreteau, Salem et al., 2007; Kalluri, Gilruth et al., 2007); and
- challenges in developing methods to identify the appropriate data resolutions and to integrate these

data into spatial units that are relevant to disease transmission (Rogers & Randolph, 2003). Another issue identified is related to validity and veracity of maps produced. Risk map is the outcome of modelling disease transmission based on spatial and temporal data (Kitron, 2000). It attempts to represent ecological, biological, behavioural processes influenced by host, pathogen and the surrounding environment (Karesh, Dobson et al., 2012; Slingenbergh, Gilbert et al., 2004; Wilcox & Gubler, 2005). Such a process may misrepresent the dynamic process for any given disease if, for example, (i) the input data are outdated, or simply wrong; (ii) the explanatory variables being used are selected wrongly; (iii) the epidemiological assumption underlying the risk map is wrong (Hirzel, Hausser et al., 2002; Rogers, 2006; Rogers & Randolph, 2003; Wood, Beck et al., 1991; Woolhouse, 2011). Besides, a pathogen or a disease will not occupy all suitable habitats. Hence even with the power combining GIS, RS, computer technology and statistical tools, risk maps developed by using disease risk factors and their surrogates, individually or collectively could only estimate the likelihood of disease introduction or spread at a geographic location and a time period. To improve the predictability of any risk map, validation with disease presence and absence data is imperative. This can be done through ground truthing, facilitated by checking the agreement between the predictions and observations or historical records by statistical methods such Area Under Curve (AUC) of Receiver Operating Characteristic (ROC) and Kappa statistic (Brownstein, Skelly et al., 2005; Glass, Cheek et al., 2000; Paul, Held, & Toschke, 2008; Sumption, Rweyemamu, & Wint, 2008; Yang, Vounatsou et al., 2005). Besides, constant updating data of the risk maps is essential.

Up to now, a majority of the disease risk maps developed have been static ones. Such a situation may be changed in the near future. With the availability of big data and improvement in GIS and RS techniques, Hay, et al envisioned that evolving or even real time disease risk maps could become a reality (Hay, George et al., 2013). Big Data is a term used to describe information assemblages having either big volume, high frequency of update (velocity) or diversity (variety) (Najjar, 2014). Big data can be health outcomes or information on environmental, climatological descriptors generated from satellites.

Systematic combination of risk mapping and prospective surveillance encompassing epidemiological, environmental and socio-economic data is a new concept which has only been sporadically investigated and deserves greater attention in the future (Aagaard-Hansen, Sorensen, & Chaignat, 2009).

Sentinel surveillance

Sentinel surveillance is one form of surveillance in which activities focus on specific subpopulations to enhance detection of disease and/or improve the cost-effectiveness of surveillance (McCluskey, 2003b).

Sentinel surveillance can be defined as a risk-based approach because sentinels are: (i) more likely to be exposed to a disease pathogen; (ii) more likely to be highly susceptible or vulnerable to the disease/vector and (iii) the disease is preferably more detectable in sentinels than in other susceptible

species (McCluskey, 2003a; Racloz, Griot, & Stärk, 2006). For instance, the most at risk populations such as injecting drug users, female sex workers as well pregnant women were recruited as the sentinels for HIV surveillance to understand the disease spread (Celentano, Akarasewi et al., 1994; Chen, Wang et al., 2012; Kigadye, Klokke et al., 1993). School age students, as sentinels, were monitored for human influenza outbreak and other emerging infectious diseases (Lenaway & Ambler, 1995; Soh, Cook et al., 2012). Sentinel herds in high risk areas were used for surveillance on infectious diseases, like bluetongue disease and birds for West Nile Virus surveillance, for early detection of these disease incursion (Carney, Ahearn et al., 2011; Chaintoutis, Dovas et al., 2015; Komar, 2001; Mostashari, Kulldorff et al., 2003; Racloz, Griot & Stärk et al., 2006; Roberts & Foppa, 2006). Sentinels were also used to study epidemiology of disease (McCluskey, 2003b). Besides, through modelling work, Smieszek and Salathe suggested simple proxies such as collocation ranking method, based on information collected via wireless wearable sensors, which may effectively identify subpopulations suitable as sentinels for human influenza early warning and surveillance (Smieszek and Salathé 2013).

Apart from being used for monitoring the spread of existing diseases in terms of direction, scale and changes in the prevalence or incidence (L'Herminez & Mbizvo, 1997; Reintjes & Wiessing, 2007; Richard, Vidondo, & Mäusezahl, 2008), environmental health hazards (Rabinowitz, Peter et al., 2005; Van der Schalie, Gardner et al., 1999), antibiotic and antiviral resistance (Schwarcz, Zenilman et al., 1990), sentinel surveillance has also been employed for evaluating efficacy of disease mitigation strategies such as vaccination programs (Janjua, Skowronski et al., 2012; Skowronski, Janjua et al., 2013; Skowronski, De Serres et al., 2009; Suarez, 2005) and as an early warning system for emergence of disease (Kulasekera, Kramer et al., 2001; McCluskey, 2003b; Snow, Newson et al., 2007). Specific cohort (can be human, animals and disease vectors) in pre-defined location within a geographic area, such as farm and health facility have been used as sentinels in surveillance systems (Celentano, Akarasewi et al., 1994; Doherr, Heim et al., 2001; Komar, 2001; Soh, Cook et al., 2012).

A framework for assessing the utility of potential animal sentinels for surveillance purpose was developed by Halliday, et al. The framework comprises the following three components:

- the sentinel response to the pathogen;
- the relationship between sentinel and target populations and
- routes of transmission to both target and sentinel populations.

Further details of the framework are illustrated by Figure 2-9. Using HPAI H5N1 as an example, they demonstrated how to use the framework to identify domestic chickens and ducks as the sentinels in a country with underdeveloped disease surveillance and reporting structure (Halliday, Meredith et al., 2007). The process was transparent and informed by the existing evidence.
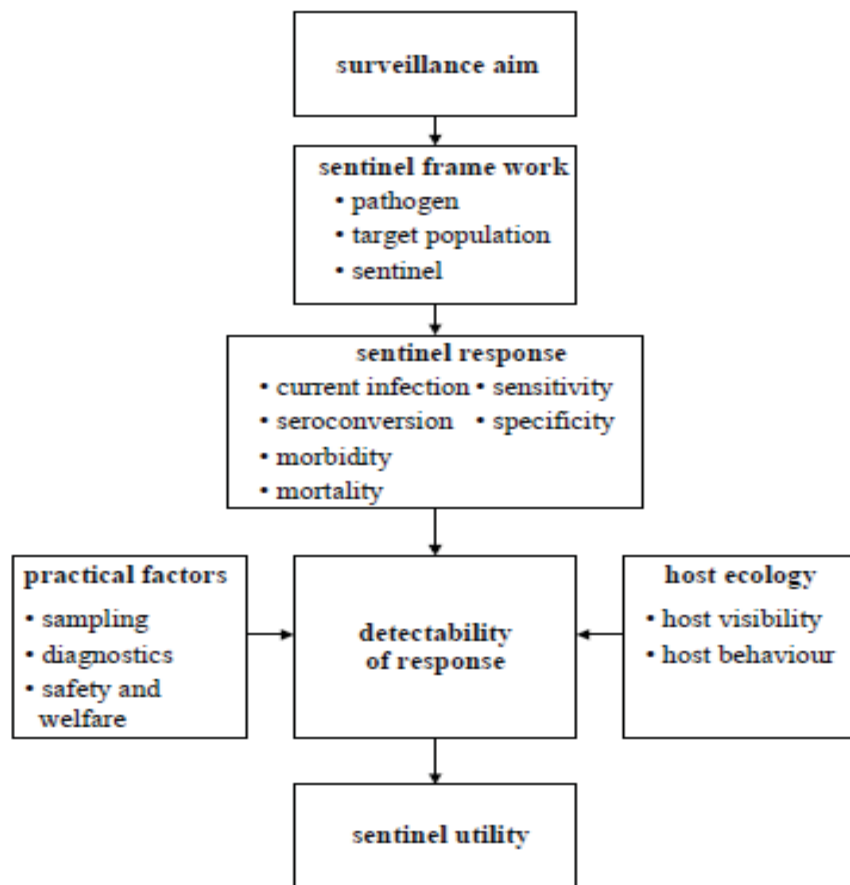
Figure 2-9. The Sentinel Framework in Context
Source: Halliday, J. E., A. L. Meredith, et al. (2007). "A framework for evaluating animals as sentinels for infectious disease surveillance." Journal of the Royal Society Interface 4(16): 973-984.

Other explanations on sentinel selection include:

- the pathogen must be a known one, i.e. animal sentinels cannot provide the solution to the question of how to carry out surveillance for pathogens that are currently unknown;

- minimally, a sentinel and its represented target population must be spatial associated. Besides, the sentinel and target population may also be epidemiologically linked such that the sentinel may act as a source of infection for the target population, as is the case with arthropod vector surveillance;

- early warning sentinels are those used to provide a predictive signal of risk to the target population. In most cases, early warning sentinels are highly visible and develop a very obvious response to the pathogen. The disease sets on earlier in sentinels than in a target population. In addition, data provided by sentinels with these qualities can be more rapidly processed, analysed and acted upon; and

- sentinels can also be used retrospectively to provide evidence of the timing of pathogen introduction and spread through a target population (Halliday, Meredith et al., 2007).

Other approaches

Benschop, et al. used time series analysis methods to identify seasonality or other temporal autocorrelation of Salmonella sero-prevalence data collected over more than 10 years. Lagged scatter-plots, autocorrelation (ACF) plots and partial autocorrelation (PACF) plots were used to identify temporal autocorrelation in the aggregated weekly data, and an autoregressive integrated moving-average (ARIMA) process was used for predictive modelling. The study confirmed a declining trend in Salmonella sero-prevalence over the years, no seasonality in sero-prevalence, and no need for more frequent sampling at the farm level at intervals of less than every 10 weeks (Benschop, Stevenson et al., 2008). In another study, using mixed effects logistic regression model, Benschop et al. identified west of Denmark experienced the highest risk for Salmonella (Benschop, Spencer et al., 2010). She and others further developed a zero-inflated binomial model to predict which farms were most at risk. They concluded an improved risk-based surveillance strategy informed by the model, though less sensitive, could result in significant cost savings (Benschop, Spencer et al., 2010).

There exist a vast and growing suite of methods for outbreak detection and identification of temporal-spatial aberrations of diseases occurrence. The methods can be classified into either testing-based or model-based approaches. A list of criteria have been proposed by Robertson, et al. for selection of a suitable method (Table 2-2) (Robertson, Nelson et al., 2010).

Table 2-2. Contextual factors for evaluation of methods for space-time analysis of disease surveillance

Factor	Description
Scale	The spatial and temporal extent (e.g., local, national, regional, global)
Scope	The intended target of the system (e.g., single disease/multiple disease, single host/multiple host, known pathogens/unknown pathogens)
Function	The objective(s) of the systems (outbreak detection, outbreak characterization, outbreak control, case detection, situational awareness, biosecurity and preparedness
Disease characteristics	Is the pathogen infectious? Is this a chronic disease? How does it spread? What is known about the epidemiology of the pathogen?
Technical	The level of technological sophistication in the design of the system and its users (data type and quality, algorithm performance, computing infrastructure and/or reliability, user expertise)

Source: Robertson, Colin, et al. "Review of methods for space–time disease surveillance." Spatial and Spatio-temporal Epidemiology 1.2 (2010): 105-116.

2.3.3.3. Risk-based requirement

Risk-based requirement is defined as using prior or additional information about the probability of hazard occurrence to revise the surveillance intensity required to achieve the stated surveillance purpose (Hoinville, Alban et al. 2013).

A number of approaches for risk-based requirement were identified by this review. They have been used either for demonstration of freedom from a disease or estimation of prevalence of disease, or disease outbreak early warning. The statistical theories behind the design and analyses of these methods are different (Cameron & Baldock, 1998; Heckathorn, 1997; Martin, Cameron, et al., 2007a; Prattley, Morris et al., 2007; Raymond, Ick, et al., 2007).

Scenario tree modelling

In animal health, a significant proportion of risk-based surveillance practices have been focused on demonstration of disease freedom although it can also be used for disease early warning and monitoring disease prevalence (Cameron, 2012). For risk-based approaches, scenario tree modelling approach is often used to estimate sample size and surveillance system sensitivity. A typical scenario tree model is an inverted tree structure that visualizes the surveillance system components (SSC). It also details all steps along the surveillance process through "nodes, limbs/branch and direction" and the interrelationships of all factors affecting a given SSC outcome (Figure 2-10) (Martin, Cameron et al., 2007a; Martin, Cameron, & Greiner, 2007b).
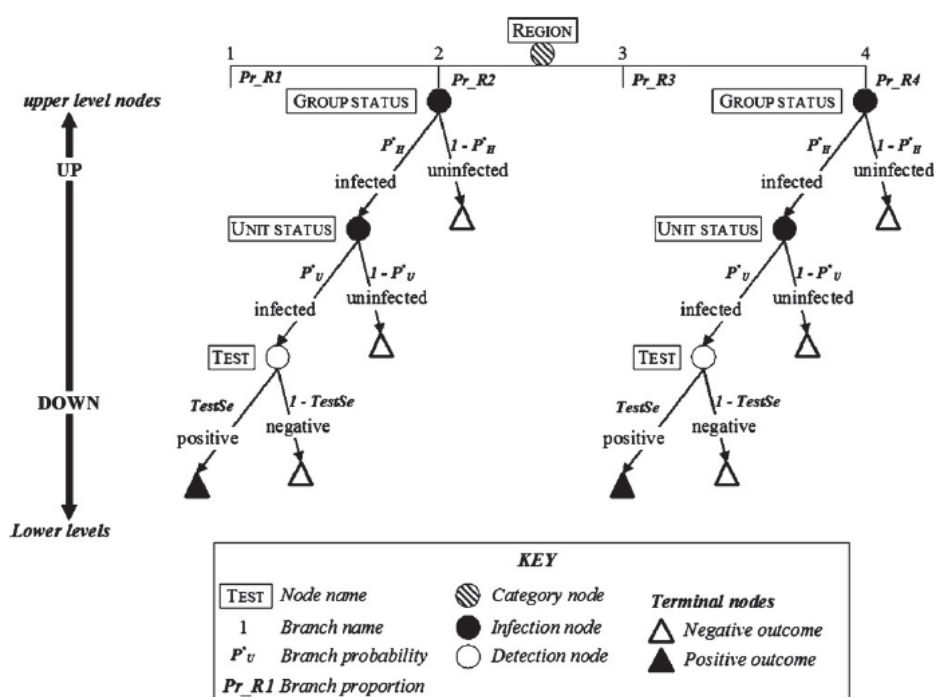
Figure 2-10. Stylized scenario tree
Source: Martin, P. A. J., A. R. Cameron, and M. Greiner. "Demonstrating freedom from disease using multiple complex data sources: 1: A new methodology based on scenario trees." Preventive Veterinary Medicine 79.2 (2007): 71-97.

The methodology of stochastic scenario tree modelling for facilitating implementation of risk-based surveillance to demonstrate freedom from disease was initially introduced by Martin and his colleagues. Methods on where to find the information for parameterizing the needed inputs the model and the formula for estimating sensitivity of SSC and the sensitivity of whole system, probability of disease

freedom were proposed (Martin, Cameron et al., 2007a; Martin, Cameron & Greiner et al., 2007b). Using this methodology, Martin et al. demonstrated probabilities that the population was free from CSF at each of the design prevalence of 0.001, 0.005 and 0.01, after a year of accumulated negative surveillance data, were 0.91, 1.00 and 1.00; targeting adults and herds from South Jutland was estimated to give approximately 1.9, 1.6 and 1.4 times higher the surveillance sensitivity that that of a proportionally representative sampling program for three among-herd design prevalence levels.

Building upon the works by Martin and others, to standardize the approach of scenario tree modelling for substantiating freedom from disease, Vanderstichel et al. proposed an approach with suggestions on the defined inputs, outputs and validation methods for implementing scenario tree modelling (Figure 2-11). They then used the approach to demonstrate market hogs in Canada was free from Trichinella. The team claimed that a standardized approach could increase transparency in comparing countries freedom claims based on the model output such as probability of freedom. It would also be helpful when comparing different studies to identify similarities and differences between them (Vanderstichel, Christensen et al. 2013). This approach was adopted by others (Christensen, El Allaki, & Vallières, 2014).
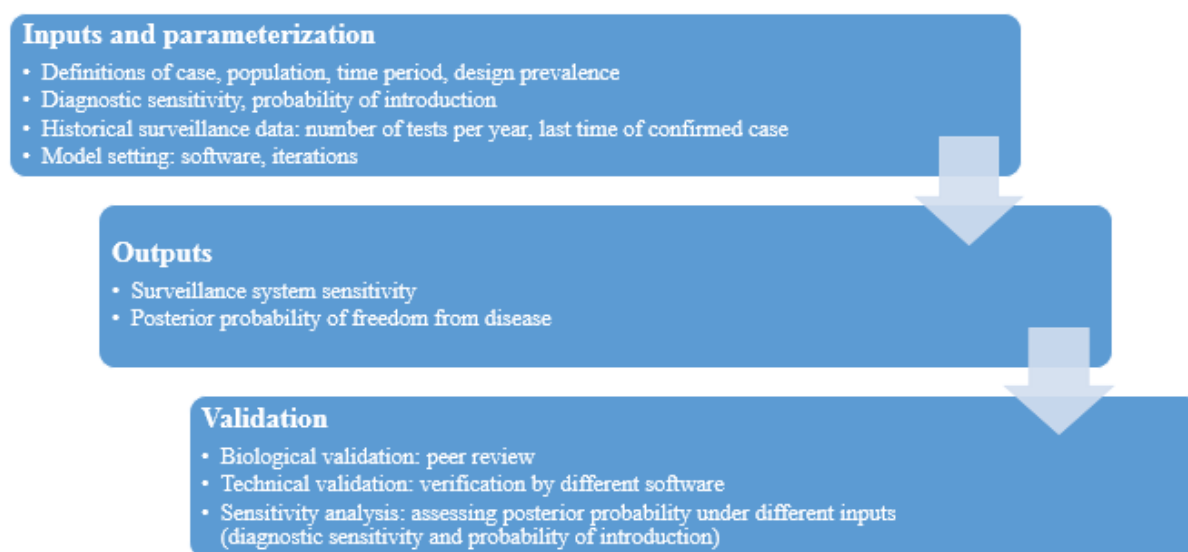
Figure 2-11. Standardized approach for demonstration of freedom from disease by using scenario tree modelling
Source: adapted from Vanderstichel, Raphaël, et al. "Standards for reporting surveillance information in freedom from infection models by example of Trichinella in Canadian market hogs." *Preventive veterinary medicine* 111.1 (2013): 176-180.

A risk-based surveillance program was designed for Trichinella in Denmark by Alban, et al. The program targets all out-door reared pigs as well as all sows and boars. Using scenario tree modelling, Alban, et al. simulated the effect of implementing this risk-based surveillance. They concluded that the RBS could reduce the total number of pigs surveyed from 23 million to 610,000 a year while the

probability of freedom from Trichinella in Denmark remains as high as 95% and above even under the worst case scenario of low surveillance system sensitivity (Alban, Boes et al., 2008).

Blickenstorfer et al. confirmed that the surveillance approach combining a surveillance component by random sampling strategy with a risk-based surveillance component was more effective in demonstrating freedom from two diseases respectively than pure random sampling approach (Blickenstorfer, Schwermer et al., 2011). Since risk factor for any given farms could change over time, they recommended risk factors and their relative risks be reviewed in regular time intervals. The similar suggestion was also made by others (Christensen, El Allaki et al. 2014).

Key assumptions for implementation of scenario tree modelling were summarized as follows: (i) all final results (i.e. after completion of any diagnostic follow-up) from the surveillance system are consistent with country or zone freedom from disease; (ii) specificity of surveillance system is 100 percent; (iii) all units under the same surveillance system component (SSC) need to be considered independently of each other with regard to probability of being infected hence the units processed under a given surveillance system component are representative of the population (Martin, Cameron et al., 2007a).

A number of online software programs have been developed for the implementation of scenario tree modelling. These at least include the one developed by Sergeant, et al. (Sergeant, Cameron, et al., 2009) and the one developed by AusVet[1].

The review has revealed that implementation of scenario tree modelling is a data intensive process, heavily relying on historical data which might be challenging for certain diseases or data paucity countries. For some inputs, estimation has to rely on experts' opinion (Alban, Boes et al., 2008; Christensen, Stryhn et al., 2011; Martin, Cameron et al., 2007a).

Discounting historical evidence
One particular interest for risk-based surveillance community is on how to use historical information from repeated surveys and information on risk related to importation to refine risk-based surveillance. The rationale behind such thinking is that those information should be valued and taken into consideration to refine confidence on freedom from disease and sample size for demonstration of freedom from disease. All the discussions around this topic can be traced back to the early works by Cannon (Cannon, 2001).

In a study, Hardon et al. detailed a five steps' procedure for estimation of probability of freedom from a disease based upon the historical information and calculation of sample size: (i) estimation of the

---

[1] http://freedom.ausvet.com.au/content.php?page=build

probability of disease freedom from the previous survey; (ii) estimation of probability of disease introduction though animal importation for the time period between the previous and the follow up surveys through risk assessment; (iii) calculation of adjusted probability of disease free; (iv) calculation of the required confidence level for the follow up survey; and (v) calculation of the required sample size for the follow up survey. They then used the methodology to estimate the sample size for the surveillance on enzootic bovine leucosis (EBL) and *Brucella melitensis* in sheep and in goats respectively in Switzerland. They concluded that the sample size for the documentation of freedom from EBL and *Brucella melitensis* in sheep and in goats could be reduced from 2,325 to 415 cattle herds, from 2,325 to 838 sheep herds and from 1,975 to 761 goat herds respectively (Hadorn, Rufenacht, et al., 2002). The formulae for calculating adjusted probability of disease freedom is illustrated below, where PFPS stands for probability of disease freedom in the previous survey, m stands the total number of countries exporting animals to Switzerland, p(imp-) is the probability of no infected animals is imported (defined by not to exceed the threshold for disease freedom) at the time of current survey.

$$PFPSadj = PFPS \times \prod_{i=1}^{m} P(imp-)i$$

Martin et al. proposed a method called temporal discounting of past surveillance data. He and the colleagues proposed to use a Bayes' formulae to update the prior estimation of the confidence that the population is not infected with the new evidence (see the formulae below, PriorPinftp stands for probability that the population was infected in the previous time period; SSetp, surveillance system sensitivity for the current time period; PostFfreetp, the adjusted probability of free from disease) (Martin, Cameron et al., 2007a; Martin, Cameron et al., 2007b). This approach has been widely used for demonstration of disease freedom for various diseases and in different countries (Christensen, Stryhn et al., 2011; Frossling, Agren, et al., 2009; Goutard, Roger et al., 2007; Flavie L. Goutard et al., 2012; Murphy, Wahlström et al., 2012; Wahlström, Frössling et al., 2010; Welby, Meroc et al., 2013).

$$PostFfreetp = \frac{1 - PriorPinftp}{1 - PriorPinftp \; X \; SSetp}$$

Schwermer et al. combined the two above approaches by taking into consideration of loss of confidence due to reduced time value of historic information and loss of confidence due to import risk and developed a method for estimation of probability of freedom from a disease. This approach results in further reduction in sample size for demonstration of freedom from disease (Schwermer, Reding, & Hadorn, 2009).

Alban et al. proposed a standardized risk-based surveillance strategy for Trichinella surveillance in Europe in order to reduce the cost spent on the Trichinella surveillance programs in Europe. Based upon risk analysis, they categorized the member states into three classes according to historical epidemiological status. Farm type, species and number of samples were proposed accordingly for surveillance purpose for each of the three classes (Alban, Pozio et al., 2011) (Table 2-3). One of the rationales behind this framework is that in case historical data are used, the sample size can be reduced while probability of freedom from disease would remain at the similar level.

Table 2-3. Suggested framework for simplified sampling scheme for Trichinella in European Union

Class	Farm Type	Species and sampling
1	• Controlled housing • Non-controlled housing	• All Trichinella-susceptible animals destined for human consumption; • All sows and boars, horse farmed • Wildlife testing optional unless meat for human consumption
2	• Controlled housing • Non-controlled housing	• Fattening pigs: proportionate sampling for fattening pigs to demonstrate surveillance sensitivity of ≥95% to detect infection of ≥1 case/million in fattening pigs for controlled housing; all fattening pigs for non-controlled housing • All sows and boars, horse farmed or hunted wild boar • Other wildlife: mandatory if move into class 3b (all pigs) anticipated, otherwise optional
3	• Non-controlled housing	• Fattening pigs • All sows and boars, horse farmed or hunted wild boar • Other wildlife: optional or proportion to demonstrate a low level in wildlife <0.1%

Source: Alban, L., et al. "Towards a standardized surveillance for Trichinella in the European Union." Preventive veterinary medicine 99.2 (2011): 148-160.

Respondent driven sampling and Time-location sampling (TLS)

These two methods are probability sampling approaches that have been increasingly used for investigation and surveillance in hard to reach or hidden populations for HIV/AIDS.

Respondent-driven samplng (RDS) is a relatively new sampling and sample size calculation method. It was firstly used for studying HIV-related risk behaviors among injecting drug users and then expanded to surveillance among other most at-risk populations for HIV (Heckathorn, 1997; Wattana, van

Griensven et al., 2007). Subjects of RDS are recruited via snowballing, which means the current sample members recruit future sample members. Sampling process starts with the selection of a set of people (seeds) in the target population. After participating in the study, these seeds are each provided with a fixed number of unique recruitment coupons that they use to recruit other people they know in the target population. After participating in the study, these new sample members again are provided with recruitment coupons for recruiting others. The sampling continues in this way, with subjects recruiting more subjects, until the desired sample size is reached. Through RDS, unbiased estimates of the prevalence of certain traits in these populations are able to be achieved (Volz & Heckathorn, 2008). A RDS has to meet four criteria: (i) documentation of who recruited whom must be tracked, generally through a coupon system; (ii) recruitment must be rationed with generally no more than three coupons allotted per 'seed', (iii) information on personal network size must be gathered and recorded; and (iv) recruiters and recruits must know one another (i.e. have a preexisting relationship)(Magnani, Sabin et al., 2005).

Software program and user manual for RDS implementation have been developed to facilitate utilization of this methodology (Spiller, Cameron, & Heckathorn, 2012). So far, RDS has only been used for HIV/AIDS related studies. Despite many years' use and having become a gold standard for HIV surveillance among most at-risk populations (MARPs), further standardization in utilization seems needed (Malekinejad, Johnston et al., 2008).

In human health sector, TLS has been used to investigate target populations when they congregate at certain physical or virtual locations (Ferreira, De Oliveira et al., 2008; Magnani, Sabin et al., 2005; Wei, McFarland et al., 2012). As locations may structure social and sexual networks, data on these locations may help to rule out particular network structures and estimate population composition (Karon & Wejnert, 2012). While only a fraction of venues are sampled, it may be possible to get indirect insights into other venues through reports by individuals who visit multiple venues. When TLS is implemented, sites for congregation are enumerated in a preliminary ethnographic mapping or pre-surveillance assessment exercise; the list of sites so developed is used as a sampling frame from which a probability sample of sites will be chosen. Data are gathered from either all or a sample of subgroup members found at the site during a pre-defined time interval Two pre-conditions for a sound TLS are: (i) all or a very high percentage of sites where subgroup members congregate are identified so that they can be included in the sampling frame, and (ii) all or a very high percentage of subgroup members visit such sites at least periodically. However such conditions are not always satisfied hence TLS suffers from potentially unacceptable levels of bias (Magnani, Sabin et al., 2005).

In one study, Wei, et al compared the two methods for studying black men who have sex with men. They found that prevalence of HIV and unrecognized infections were slightly higher among RDS participants who were less likely to have a main partner, but more likely to have a female partner and

have both male and female partners, and reported greater methamphetamine, crack and heroin use. In other words, RDS reached the riskier subgroups of this population (Wei, McFarland et al., 2012).

Portfolio theory

Prattley et al. demonstrated how to use portfolio theory to allocating number of samples to be tested by location and by temporal duration based upon the comprehensive risk score per location and per period and the associated uncertainty. The study illustrated that both historical data and expert opinion could be incorporated into the risk assessments that should be updated from time to time, so that a dynamic risk landscape could be constructed to map the risk of disease as it may change over time across the area of concern (Prattley, Morris et al., 2007).

Weighted Surveillance Approach

Walsh, et al used a weighted approach for Chronic Wasting Disease (CWD) surveillance in USA. The animals under surveillance were grouped into eight demographic strata different in age, sex, etc. For calculating sample size for each stratum, a weight system was used. Strata with higher CWD prevalence and low inclusion probability receive higher weights, which meant a higher number of individuals would be sampled from them. The weight of a stratum is actually the risk ratio or odd ratio between it and the stratum with the lowest CWD prevalence. The authors simulated the effects of increase proportion of samples collected from high risk strata and concluded that by implementing the weighted surveillance system, fewer samples would need to be collected and examined while maintaining or improving current surveillance standards (in terms of time needed for detection of first CWD) (Walsh & Miller, 2010).

Other relevant discussions on risk-based requirement

Cameron discussed when to use the comprehensive concept of risk, in terms of likelihood of introduction, exposure, and consequence of exposure account in RBS approaches. His suggestions are summarized in Table 2-4.

Table 2-4. Summary on when likelihood and consequence need to be used for different RBS

Type of RBS	likelihood	consequence	
		Freedom from disease	Early detection
Risk prioritization	Yes	Yes	Yes
Risk sampling	Yes	No	Yes
Sampling within stratum	No	No	No

Note: RBS stands for risk-based surveillance
Source: Cameron, A. R. "The consequences of risk-based surveillance: Developing output-based standards for surveillance to demonstrate freedom from disease." Preventive veterinary medicine 105.4 (2012): 280-286.

The review found that many risk-based sampling based upon scenario tree models tend to agree on that although individuals from different risk stratum should be sampled in

correspondence to their risk levels, individuals within each risk stratum sampled randomly since risk level within the risk stratum is assumed the same (Cameron, 2012). Although such an assumption help simplifying sampling process, the rationale for such an approximation has not been discussed.

Diagnostic test performance has been assumed to be perfect in most studies using scenario tree modelling methodology. Such an assumption may not be true in the real world. To address imperfect test, a simple modification of the approximate formulae considering test sensitivity has been developed by MacDiarmid based upon the approximation method of Cannon and Roe (Cannon & Roe, 1982; MacDiarmid, 1988). In the formulae below, *P* denotes probability of *x* positive sample; *n,* sample size; *p,* designed prevalence; *Se,* test sensitivity; *Sp,* test specificity.

$$P = \binom{n}{x}[pSe + (1 - p)(1 - Sp)]^x[p(1 - Se) + (1 - p)Sp]^{n-x}$$

2.3.3.4. Risk-based analysis

The definition of risk-based analysis is using prior or additional information about the probability of hazard occurrence (including contextual information and prior likelihood of disease), to revise conclusions about disease status (Hoinville, Alban et al., 2013).

Gustafson, et al developed a decision framework for estimating disease probability by combining evidence streams. This evidence aggregation model is based on the odds form of Bayes' theorem. In the study, disease risk score elicited from experts, used for estimating the predicted occurrence of risk factors among the disease affected versus unaffected watersheds, was combined with surveillance data to produce a risk adjusted posterior probability of the disease for a given watersheds. The method provides a flexible framework for iterative revision of disease freedom status as knowledge and data evolve (Gustafson, Klotins et al., 2010).

2.4. Discussion: key observations on RBS

2.4.1. Key assumptions for RBS

The underlying principle of RBS is that by focusing surveillance resources on sub-populations in which disease is more likely to occur, the cost of obtaining the required information will be reduced, or the value of the information will be increased, or both. There is also an implicit assumption that there will be little or no reduction in effectiveness of surveillance if this approach replaces part or all of the prior investment in classical surveillance techniques. RBS is only effective if these assumptions are

sufficiently met. Failure to meet important assumptions may result in inaccurate identification of risk strata, which could reduce the technical and economic benefits of RBS. The key assumptions or prerequisites for each individual methodology have been documented in the previous sections.

2.4.2. Benefits from RBS

Benefits from implementation of risk-based approaches identified from the review include: (i) improvement in technical efficiency of the surveillance system; (ii) improving economic efficiency and (iii) improved feasibility of surveillance program implementation.

Improvement in technical efficiency

Compared to conventional surveillance using probability sampling, pure or even partial risk-based approaches could: reduce the number of samples to be tested while maintaining the same level of confidence, improve the probability of providing sufficient evidence for freedom from a disease, shorten the time for disease outbreak detection and improve overall surveillance system sensitivity, or sensitivity ratio (Table 2-5).

Table 2-5. Summary on technical efficiency improvement by RBS

TE* measurement	Some of selected references
Sample size	(Blickenstorfer, Schwerme et al., 2011; Hadorn, Rufenacht et al., 2002; Presi, Staerk et al., 2008; Reist, Jemmi & Staerk, 2012; Schwermer, Reding et al., 2009; Walsh & Miller, 2010; Willeberg, Nielsen, & Salman, 2012)
Probability of disease free	(Alba, Casal et al., 2010; Hadorn, Racloz et al., 2009; Reist, Jemmi & Staerk, 2012; Tavornpanich, Gardner et al., 2006)
Time for disease detection	(Carney, Ahearn et al., 2011; Chaintoutis, Dovas et al., 2015; Healy Reisen et al., 2015; Hii, Rocklov et al., 2012; Komar, 2001; Kulasekera, Kramer et al., 2001; Mostashari, Kulldorff et al., 2003; Racloz, Griot & Staerk et al., 2006; Roberts & Foppa, 2006)
Surveillance system sensitivity, sensitivity ratio**	(Alba, Casal et al., 2010; Calvo-Artavia, Nielsen, & Alban, 2013; Flavie L. Goutard et al., 2012; Knight-Jones, Hauser et al., 2010)

Note: *TE, technical efficiency; ** Sensitivity ratio means surveillance system component sensitivity divided by the reference surveillance component sensitivity.

Improvement in economic efficiency

RBS approaches outperform conventional surveillance approaches by lesser surveillance cost, higher cost effectiveness, lower cost effectiveness ratio, or higher effectiveness ratio (Table 2-6).

Table 2-6. Summary on economic efficiency improvement by RBS

EE* measurement	Some selected reference
Reduced surveillance cost, cost saving**	(Reist, Jemmi & Staerk, 2012; Tavornpanich, Gardner et al., 2006; Walsh & Miller, 2010)
Lower cost-effectiveness ratio [1]	(Knight-Jones, Hauser et al., 2010)
Higher cost-effectiveness [2]	(Healy, Reisen et al., 2015)
Effectiveness ratio[3]	(Calvo-Artavia, Nielson & Alban, 2013)
Net economic effect	(Calvo-Artavia, Nielson & Alban., 2013)

Note: *EE, economic efficiency; **compared to the cost of conventional surveillance by using random sampling; 1, cost divided by surveillance component sensitivity; 2, number of positive results per US$1,000 spent; 3, net economic effect divided by change of surveillance system sensitivity.

Improved feasibility of surveillance implementation

As mentioned earlier, RBS can significantly reduce the number of samples to be collected, which would reduce complexity and burden in implementing surveillance strategies. Again, in the study on bovine cysticercosis surveillance in Denmark, Calvo-Artavia et al. illustrated that a gender (risk factor) based sampling approach would be easier and more accurate for implementation compared to using other risk factors while the level of technical performance is comparable to those using other risk-based surveillance scenarios (Calvo-Artavia, Nielson & Alban, 2013). Using animal sentinels such as dead birds for West Nile virus disease surveillance is operationally less challenging than other surveillance approaches (Komar, 2001).

2.4.3. Challenges related to RBS

Political hurdles

Probability for detection of positive samples is possibly higher by using risk-based surveillance approaches than that by using non-risk-based surveillance. This may lead to trade restrictions and other negative consequences (Stark, Regula et al., 2006; Tsai, Scott et al., 2009). In addition, implementation of risk-based surveillance does not mean it is inexpensive. For instance, social network approach for identifying individuals, sub-populations and potential hotspots can be highly resource and labor intensive (Bolton, McCaw et al., 2012; Smieszek & Salathé, 2013). However setting incentives can encourage RBS. For instance, EU granted regions or countries having a negligible risk of Trichinella in domestic swine a reduced surveillance and testing approach for the disease in swine carcasses (Alban Boes et al., 2008; Alban, Pozio et al., 2011). Furthermore, RBS can be technically sophisticated and might not be intuitively straightforward to decision makers, thus communication on the results to them and winning their support could be challenging (Reist, Jemmi & Staerk, 2012).

To practice risk-based approaches for surveillance requires a mindset change and harmonization of surveillance approaches across countries (L. Alban et al., 2011; Pozio et al., 2010). Countries or a region need to set up a proper legal framework to allow focusing on outcome based standards for

surveillance (Martin, Cameron & Greiner, 2007b) and allow flexibility in surveillance system design (Cameron, 2012).

Validation and veracity

A variety of methods have been implemented to make sure RBS approaches are valid as well as accurate. For instance, sensitivity analysis has been widely used to assess the robustness of the findings from a given RBS approach. Spearman rank correlation coefficient was calculated to check the agreement between two rounds of assessment results given by the assessors participating in MADM. Cohen Kappa and AUC, etc. were used for assessing the agreement between expected and observed risk surfaces. Veracity of risk maps was recommended to be checked by ground-truthing. Diseases evolve over time, so do the risk factors, this necessitates the inputs and outcomes of RBS approaches to be reviewed on regular basis.

Standardization

The review has identified standardization efforts for: MADM for risk prioritization, social network analysis, respondent driven sampling, scenario tree modelling, as well disease risk mapping. Standardization takes forms of development and promotion of a common framework, guidelines as well as development of software programs. The arguments for standardization include transparency and easier cross comparison so that similarities and differences can be identified and discussed. However, one comment on standardization cautions people that the desire to harmonize surveillance programs via standardization between countries can impede the evolution of efficient risk-based surveillance strategies tailored to national risks (Stark, Regula et al., 2006).

Data issues

RBS can be data intensive. The design and implementation of a risk-based approach for surveillance requires: quality baseline information, often years of prior data on difference in occurrence of disease between population strata and risk factors and their influence over disease occurrence (Alba, Casal et al., 2010; Carroll, Au et al., 2014; Christensen, El Allaki et al., 2014; Hadorn, Racloz et al., 2009). These entail that information on disease, hosts, risk factors/drivers have to be generated, updated and curated. All these challenges exacerbate the issue of data scarcity in developing countries. For instance, few developing countries have an animal movement registry. While multi-temporal satellite data are available for an extended time period, the availability of georeferenced and spatially explicit disease data for the same temporal period is still less common, especially in developing countries. Generating data often requires good research and technical capacities (Paul, Held et al., 2008). These challenges may explain partially why a majority of risk-based surveillance activities have been implemented in developed economies.

Risk stratification based upon historical data may only represent historical reality. Caution has to be applied since risk factors may change overtime especially in the case of detecting emerging infectious

diseases, which requires constant re-evaluation of the assumptions underlying the risk-based surveillance approach. For instance, animal movement has to be generated during or right before disease incursion and spread, or in another words, in an epidemiological relevant time window (Benschop, Stevenson et al., 2008; Ortiz-Pelaez & Pfeiffer, 2008).

Even though data exist, data ownership and accessibility may also present a challenge between different sectors, geographic areas within a country and particularly between countries. Political support, proper governance framework, comparability of data and information infrastructure are all in play to support effective data sharing and access (Weinberg, Waterman et al., 2003).

Availability of remotely sensed data has been improved greatly (Hay, Tatem et al., 2006). However transformation or conversion of the remotely sensed data on environmental variables is essential to convert the raw data into epidemiologically meaningful values. Besides, such datasets should be made available to epidemiologists in real-time (if possible automation) and in a format that they can readily use as inputs for their modelling (Eisen & Eisen, 2011; Hay, Battle et al., 2013).

Certain data that are useful for RBS are sensitive and private. For instance, questions about sexual mixing behavior are inherently personal, and questions about physical contacts may be considered intrusive in some communities (Eames, Bansal et al., 2015). In disease mapping, a common consideration cited is how to protect privacy of public health data (Gao, Mioc et al., 2009; Geanuracos, Weiss et al., 2007).

A number of approaches have been proposed to deal with data related issues. These include, for example: crowdsourcing (Boulos, Resch et al., 2011; Eames, Tilston et al., 2012; Prah, Copas et al., 2013; Salathé, Kazandjieva et al., 2010), experts' opinion (Presi, Staerk et al., 2008), new forms of data, and open data sources such as remote sensing and data from proximity logger and animal tracking devices.

2.4.4. Evolution in risk-based disease surveillance

Output based surveillance
Methods for design and analysis of surveillance for human and animal diseases have been continuously evolving over years. Output based surveillance such as scenario tree modelling, risk-based surveillance, etc. were introduced relatively recently (Heckathorn, 1997; Martin Cameron et al., 2007a). These techniques capture the effects of differentiating surveillance efforts within population strata with different risk levels of infection or disease occurrence. When they are correctly used, they can achieve improved technical and economic efficiency in disease detection.

It was common to see that approaches, methods and even laboratory procedures are specified in detail by technical agencies such as World Health organization, Food and Agriculture Organization for a given disease (FAO, 2006; WHO, 2004). Such an approach focusing on controlling surveillance inputs and process intended to standardize surveillance so as the capacity building and quality assurance could be implemented easier and results from different countries could be compared and understood with less difficulty. Output based surveillance approach gain its popularity in recent years because of increased attention to emerging infectious diseases and public health emergencies.

Output based surveillance approach has also undergone evolution over the years. For instance, Cameron summarized output based surveillance approaches for demonstrating freedom from infection either as part of a control program or confirmation of successful eradication into three components (Table 2-7).

Table 2-7. Development of output based surveillance for demonstration of freedom from disease

Component	Description	References
1: Sensitivity	$SSe = 1 - (1 - P \times Se)^n$, at a design prevalence, flexibility is allowed in selection of approaches different in sensitivity and sample size to achieve the same level of surveillance system sensitivity; Scenario tree modelling allows using different methods for different strata	(Martin, Cameron et al., 2007a)
2: Probability of freedom	$PostPr(free) = \dfrac{PriorPr(free)}{Prior\,Pr(free) + 1 - PriorPr(free) \times (1 - SSe)}$ allows integration of multiple sources of surveillance evidence, including historical evidence and bio security levels of herds	(Cannon, 2002)
3: Expected cost of error	The probability of error: probability of residual undetected infection or 1−Pr(free) multiplied by the consequences (expressed in monetary or other terms)	No reference (at the concept stage, has not be practiced)

Source: Cameron, A. R. "The consequences of risk-based surveillance: Developing output-based standards for surveillance to demonstrate freedom from disease." Preventive veterinary medicine 105.4 (2012): 280-286.

Embracing big data

There has been no consensual definition of big data. However, big data usually have at least one or several following elements: (i) size: the volume of the datasets is a critical factor; (ii) complexity: the structure, behaviour and permutations of the datasets is a critical factor; (iii) technologies: the tools and techniques that are used to process a sizable or complex dataset is a critical factor. There is an increasing number of systems employing big data for prediction, early warning and real time reporting of infectious disease and other forms of public health events. Examples of such systems include Global

Public Health Intelligence Network (GPHIN), ProMED-mail, Google Flu Trends, HealthMap, Argus, etc. (Blench, 2008; Brownstein & Freifeld, 2007; Brownstein, Freifeld et al., 2009; Carneiro & Mylonakis, 2009).

For the development of risk surfaces for various diseases, Hay, et al envisioned that dynamic infectious disease risk maps would be available in the future with the advent of novel online data sources, such as social media, combined with epidemiologically relevant environmental information and the advances in machine learning and the use of crowd sourcing (Hay, George et al., 2013). To transform the unstructured big data from different sources into epidemiologically meaningful information and to control noise entails better IT infrastructure, system interoperability, crowdsourcing for data analysis, collaboration among epidemiologists, computer scientists as well as GIS specialists, etc. (Hay, George et al., 2013; Rogers & Randolph, 2003; Stark, Regula et al., 2006).

One health approach

One Health is defined as the collaborative effort of multiple disciplines, working locally, nationally and globally, to reach optimal health for humans, animals, and the environment (Lewis, 2008). About sixty percent of 1,415 known human infectious diseases are zoonoses (Taylor, Latham et al., 2001). Human or animal health is the downstream event of social, economic and environmental determinants (Kawachi & Wamala, 2006; Marano, Arguin & Pappaioanouet, 2007; Marmot & Wilkinson, 2005; Patz, Daszak et al., 2004; Richmond, Elliott et al., 2005; Zinsstag, Schelling et al. 2011).

The review identified more than one fourth of the publications on risk-based surveillance focused on zoonoses, an additional 3 percent on food safety and 2 percent on antimicrobial resistance. As illustrated in a previous section, surveillance on disease occurrence among sentinel animals could provide early warning or predict occurrence of the same disease among a human population (Mostashari, Farzad, et al., 2003). On the other hand, reporting of human cases often acts as sentinels for the same disease outbreaks among animals (Minh, Schauer, 2009). To understand the relationship between infectious disease occurrence and its social, environmental and economic determinants requires expertise from different disciplines and the collaboration from animal, human, environment health sectors as well as those working on wildlife.

2.5. Conclusion

The increasing tide of globalization and interconnectedness of human and animal populations throughout the world have increased the risks for infectious disease spread and the potential for significantly greater impacts (Jones, Patel et al., 2008). The key international organizations responsible for global health of human and animal populations have developed regulations, agreements and standards to mitigate these risks, such as the International Health Regulations (2005) produced by the

World Health Organization (WHO) and the Sanitary and Phytosanitary Measures (SPS Agreement) produced by the World Trade Organization (WTO) which mandate member states to improve and maintain high standards in disease vigilance and related capacities (Gostin, 2005; Reist et al., 2012). Respective standards have been developed by WHO, the Codex Alimentarius Commission (CAC), the World Organization for Animal Health (OIE) and the International Plant Protection Convention (IPPC).

Disease surveillance is an essential component in dealing with infectious diseases and surveillance programs need to be technically efficient, operationally effective and economically beneficial to ensure optimal use of the resources that are available for controlling disease across the globe.

This review concludes that:

- the value of risk-based surveillance has been exemplified by the increase in applications of the approach over time (Rodríguez-Prieto, Vicente-Rubiano et al., 2014).
- A variety of RBS approaches have been used for a wide range of animal and human health issues for demonstration of disease freedom, disease early warning, disease control, as well as understanding of disease/hazard trends.
- It is evident that compared to the conventional surveillance approaches, RBS can improve technical efficiency, reduce cost for surveillance, advance disease detection, and achieve higher net economic benefit. In some cases, RBS approaches are easier to implement.

A number of challenges have been revealed and summarized by the review. Among them, political hurdles, paucity of data, the lack of standardization/harmonization of approaches and requirements for high technical epidemiological expertise are the major ones. These challenges have limited the effective utilization of RBS in resource-poor settings.

The experiences so far in applying RBS show that there is now a need for a more integrated approach to disease surveillance - involving coordinated investigation of both animal and human populations, greater integration of different RBS approaches, increased use of new sources of data, and development of improved software tools that facilitate the design and implementation of RBS. Such developments will further enhance the global value of RBS, particularly for application in resource-poor countries.

2.6. References

Aagaard-Hansen, J., B. H. Sorensen, et al. (2009). "A comprehensive approach to risk assessment and surveillance guiding public health interventions." Tropical Medicine & International Health **14**(9): 1034-1039.

Abdullah, A. S. (2007). International Health Regulations (2005). Regional Health Forum.

Abrial, D., D. Calavas, et al. (2003). "Descriptive spatial analysis of BSE in western France." Veterinary Research **34**(6): 749-760.

Akin, S.-M., P. Martens, et al. (2015). "Climate Change and Infectious Disease Risk in Western Europe: A Survey of Dutch Expert Opinion on Adaptation Responses and Actors." International Journal of Environmental Research and Public Health **12**(8): 9726-9749.

Alba, A., J. Casal, et al. (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., J. Boes, et al. (2008). "Towards a risk-based surveillance for Trichinella spp. in Danish pig production." Preventive Veterinary Medicine **87**(3-4): 340-357.

Alban, L., E. Pozio, et al. (2011). "Towards a standardised surveillance for Trichinella in the European Union." Preventive Veterinary Medicine **99**(2-4): 148-160.

Alleweldt, F., M. Upton, et al. (2009). Cost of national prevention systems for animal diseases and zoonoses in developing and transition countries, Paris: World Organization for Animal Health (OIE).

Anderson, P. K., A. A. Cunningham, et al. (2004). "Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers." Trends in Ecology & Evolution **19**(10): 535-544.

Balabanova, Y., A. Gilsdorf, et al. (2011). "Communicable diseases prioritized for surveillance and epidemiological research: results of a standardized prioritization procedure in Germany, 2011." Plos One **6**(10): e25691.

Bansal, S., J. Read, et al. (2010). "The dynamic nature of contact networks in infectious disease epidemiology." Journal of Biological Dynamics **4**(5): 478-489.

Barbiero, V. K. (2014). "It's not Ebola… it's the systems." Global Health: Science and Practice **2**(4): 374-375.

Barrett, R., C. W. Kuzawa, et al. (1998). "Emerging and re-emerging infectious diseases: the third epidemiologic transition." Annual Review of Anthropology: 247-271.

Beck, L. R., B. M. Lobitz, et al. (2000). "Remote sensing and human health: new sensors and new opportunities." Emerging Infectious Diseases **6**(3): 217.

Benschop, J., M. Stevenson, et al. (2008). "Temporal and longitudinal analysis of Danish Swine Salmonellosis Control Programme data: implications for surveillance." <u>Epidemiology and Infection</u> **136**(11): 1511-1520.

Benschop, J., S. Spencer, et al. (2010). "Bayesian zero-inflated predictive modelling of herd-level Salmonella prevalence for risk-based surveillance." <u>Zoonoses and Public Health</u> **57 Suppl 1**: 60-70.

Berke, O. (2004). "Exploratory disease mapping: kriging the spatial risk function from regional count data." <u>International Journal of Health Geographics</u> **3**(1): 18.

Bettcher, D. W., S. Sapirie, et al. (1997). "Essential public health functions: results of the international Delphi study." <u>World health statistics quarterly. Rapport trimestriel de statistiques sanitaires mondiales</u> **51**(1): 44-54.

Blench, M. (2008). Global public health intelligence network (GPHIN). Proceedings of the conference of the American machine translation association (AMTA). Waikiki, Hawai'i: AMTA.

Blickenstorfer, S., H. Schwermer, et al. (2011). "Using scenario tree modelling for combined targeted and random herd sampling to substantiate freedom from disease." <u>Epidemiologie et Sante Animale</u>(59/60): 78-79.

Bolton, K. J., J. M. McCaw, et al. (2012). "Influence of contact definitions in assessment of the relative importance of social settings in disease transmission risk." <u>Plos One</u> **7**(2): e30893.

Boulos, M. N. K., B. Resch, et al. (2011). "Crowdsourcing, citizen sensing and sensor web technologies for public and environmental health surveillance and crisis management: trends, OGC standards and application examples." <u>International Journal of Health Geographics</u> **10**(1): 67.

Brookes, V., M. Hernández-Jover, et al. (2014). "Building a picture: prioritisation of exotic diseases for the pig industry in Australia using multi-criteria decision analysis." <u>Preventive Veterinary Medicine</u> **113**(1): 103-117.

Brookes, V. J., V. J. D. R. Vilas, et al. (2015). "Disease prioritization: what is the state of the art?" <u>Epidemiology and Infection</u> **143**(14): 2911-2922.

Brooks-Pollock, E., G. O. Roberts, et al. (2014). "A dynamic model of bovine tuberculosis spread and control in Great Britain." <u>Nature</u> **511**(7508): 228-231.

Brownstein, J. S., D. K. Skelly, et al. (2005). "Forest fragmentation predicts local scale heterogeneity of Lyme disease risk." <u>Oecologia</u> **146**(3): 469-475.

Brownstein, J. and C. Freifeld (2007). "HealthMap: the development of automated real-time internet surveillance for epidemic intelligence." Euro Surveill 12(11): E071129.

Brownstein, J. S., C. C. Freifeld and L. C. Madoff (2009). "Digital disease detection—harnessing the Web for public health surveillance." New England Journal of Medicine 360(21): 2153-2157.

Calain, P. (2007). "Exploring the international arena of global public health surveillance." <u>Health Policy and Planning</u> **22**(1): 2-12.

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

Calvo-Artavia, F. F., L. R. Nielsen, et al. (2013). "A case–control study of risk factors for bovine cysticercosis in Danish cattle herds." Zoonoses and Public Health **60**(4): 311-318.

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 F. C. Baldock (1998). "A new probability formula for surveys to substantiate freedom from disease." Preventive Veterinary Medicine **34**(1): 1-17.

Cannon, R. (2001). "Sense and sensitivity—designing surveys based on an imperfect test." Preventive Veterinary Medicine **49**(3): 141-163.

Cannon, R. (2002). "Demonstrating disease freedom—combining confidence levels." Preventive Veterinary Medicine **52**(3): 227-249.

Cannon, R. and R. T. Roe (1982). Livestock disease surveys: a field manual for veterinarians, Australian Government Publishing Service.

Cardoen, S., X. Van Huffel, et al. (2009). "Evidence-based semiquantitative methodology for prioritization of foodborne zoonoses." Foodborne Pathogens and Disease **6**(9): 1083-1096.

Carneiro, H. A. and E. Mylonakis (2009). "Google trends: a web-based tool for real-time surveillance of disease outbreaks." Clinical infectious diseases 49(10): 1557-1564.

Carney, R. M., S. C. Ahearn, et al. (2011). "Early Warning System for West Nile Virus Risk Areas, California, USA." Emerging Infectious Diseases **17**(8): 1445-1454.

Carroll, L. N., A. P. Au, et al. (2014). "Visualization and analytics tools for infectious disease epidemiology: A systematic review." Journal of Biomedical Informatics **51**: 287-298.

Cediel, N., L. C. Villamil, et al. (2013). "Setting priorities for surveillance, prevention, and control of zoonoses in Bogota, Colombia." Revista panamericana de salud publica = Pan American journal of public health **33**(5): 316-324.

Celentano, D. D., P. Akarasewi, et al. (1994). "HIV-1 infection among lower class commercial sex workers in Chiang Mai, Thailand." Aids **8**(4): 533-538.

Chaintoutis, S. C., C. I. Dovas, et al. (2015). "Surveillance and Early Warning of West Nile Virus Lineage 2 Using Backyard Chickens and Correlation to Human Neuroinvasive Cases." Zoonoses and Public Health **62**(5): 344-355.

Chen, X.-S., Q.-Q. Wang, et al. (2012). "Prevalence of syphilis infection in different tiers of female sex workers in China: implications for surveillance and interventions." BMC Infectious Diseases **12**.

Chon, H.-S., D.-G. Ohandja, et al. (2012). "A risk-based approach to prioritise catchments for diffuse metal pollution management." Science of the Total Environment **437**: 42-52.

Christensen, J., F. El Allaki, et al. (2014). "Adapting a scenario tree model for freedom from disease as surveillance progresses: the Canadian notifiable avian influenza model." Preventive Veterinary Medicine **114**(2): 132-144.

Christensen, J., H. Stryhn, et al. (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): 161-175.

Christley, R. M., G. Pinchbeck, et al. (2005). "Infection in social networks: using network analysis to identify high-risk individuals." American Journal of Epidemiology **162**(10): 1024-1031.

Ciliberti, A., D. Gavier-Widen, et al. (2015). "Prioritisation of wildlife pathogens to be targeted in European surveillance programmes: Expert-based risk analysis focus on ruminants." Preventive Veterinary Medicine **118**(4): 271-284.

Cox, R., C. W. Revie, et al. (2012). "The Use of Expert Opinion to Assess the Risk of Emergence or Re-Emergence of Infectious Diseases in Canada Associated with Climate Change." Plos One **7**(7).

Craig, M. H., B. L. Sharp, et al. (2007). "Developing a spatial-statistical model and map of historical malaria prevalence in Botswana using a staged variable selection procedure." International Journal of Health Geographics **6**.

Cringoli, G., L. Rinaldi, et al. (2005). "Disease mapping and risk assessment in veterinary parasitology: some case studies." Parassitologia **47**(1): 9-25.

Cumming, G. S., P. A. R. Hockey, et al. (2008). "Wild bird movements and avian influenza risk mapping in southern Africa." Ecology and Society **13**(2): 26-Art. 26.

Doherr, M., D. Heim, et al. (2001). "Targeted screening of high-risk cattle populations for BSE to augment mandatory reporting of clinical suspects." Preventive Veterinary Medicine **51**(1): 3-16.

Doherty, J.-A. (2000). "Establishing priorities for national communicable disease surveillance." The Canadian Journal of Infectious Diseases **11**(1): 21.

Dubé, C., C. Ribble, et al. (2009). "A Review of Network Analysis Terminology and its Application to Foot-and-Mouth Disease Modelling and Policy Development." Transboundary and Emerging Diseases **56**(3): 73-85.

Eagle, N., A. S. Pentland, et al. (2009). "Inferring friendship network structure by using mobile phone data." Proceedings of the National Academy of Sciences **106**(36): 15274-15278.

Eames, K., S. Bansal, et al. (2015). "Six challenges in measuring contact networks for use in modelling." Epidemics **10**: 72-77.

Eames, K. T. and M. J. Keeling (2002). "Modelling dynamic and network heterogeneities in the spread of sexually transmitted diseases." Proceedings of the National Academy of Sciences **99**(20): 13330-13335.

Eames, K. T., N. L. Tilston, et al. (2012). "Measured dynamic social contact patterns explain the spread of H1N1v influenza." PLoS Comput Biol **8**(3): e1002425.

East, I. J., R. M. Wicks, et al. (2013). "Use of a multi-criteria analysis framework to inform the design of risk-based general surveillance systems for animal disease in Australia." Preventive Veterinary Medicine **112**(3-4): 230-247.

Economopoulou, A., P. Kinross, et al. (2014). "Infectious diseases prioritisation for event-based surveillance at the European Union level for the 2012 Olympic and Paralympic Games." Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin **19**(15).

Eisen, L. and R. J. Eisen (2011). Using Geographic Information Systems and Decision Support Systems for the Prediction, Prevention, and Control of Vector-Borne Diseases. Annual Review of Entomology, Vol 56. M. R. Berenbaum, R. T. Carde and G. E. Robinson. **56:** 41-61.

Farine, D. R. and H. Whitehead (2015). "Constructing, conducting and interpreting animal social network analysis." Journal of Animal Ecology **84**(5): 1144-1163.

Ferreira, L. O. C., E. S. De Oliveira, et al. (2008). "Use of time-location sampling for systematic behavioral surveillance of truck drivers in Brazil." AIDS and Behavior **12**(1): 32-38.

Fevre, E. M., B. M. d. C. Bronsvoort, et al. (2006). "Animal movements and the spread of infectious diseases." Trends in Microbiology **14**(3): 125-131.

Fiebig, L. (2011). Risk-based surveillance of avian influenza in switzerland: insights into poultry farm determinants and contact networks, University_of_Basel.

FAO. (2006). "Wild bird highly pathogenic avian influenza surveillance."   Retrieved 8/14, 2015, from http://www.fao.org/docrep/010/a0960e/a0960e00.HTM.

FAO, OIE, WHO, UNICEF, UNSIC and World Bank. (2008). "Contributing to One World, One Health: A Strategic Framework for Reducing Risks of Infectious Diseases at the Animal-Human-Ecosystems Interface."   Retrieved 8/24, 2015, from ftp://ftp.fao.org/docrep/fao/011/aj137e/aj137e00.pdf.

Franks, D. W., G. D. Ruxton, et al. (2010). "Sampling animal association networks with the gambit of the group." Behavioral Ecology and Sociobiology **64**(3): 493-503.

Frossling, J., E. C. C. Agren, et al. (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.

Fuller, T., H. A. Thomassen, et al. (2011). "Using remote sensing to map the risk of human monkeypox virus in the Congo Basin." EcoHealth **8**(1): 14-25.

Gallup, J. L. and J. D. Sachs (2001). "The economic burden of malaria." The American Journal of Tropical Medicine and Hygiene **64**(1 suppl): 85-96.

Gao, S., D. Mioc, et al. (2008). "Online GIS services for mapping and sharing disease information." International Journal of Health Geographics **7**(1): 8.

Gao, S., D. Mioc, et al. (2009). "International Journal of Health Geographics." <u>International Journal of Health Geographics</u> **8**: 3.

Geanuracos, C. G., G. Weiss, et al. (2007). "Use of geographic information systems for planning HIV prevention interventions for high-risk youths." <u>American Journal of Public Health</u> **97**(11): 1974.

Gemperli, A., N. Sogoba, et al. (2006). "Mapping malaria transmission in West and Central Africa." <u>Tropical Medicine & International Health</u> **11**(7): 1032-1046.

Gilsdorf, A. and G. Krause (2011). "Prioritisation of infectious diseases in public health: feedback on the prioritisation methodology, 15 July 2008 to 15 January 2009." <u>Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin</u> **16**(18).

Gkogka, E., M. W. Reij, et al. (2011). "Risk-based Estimate of Effect of Foodborne Diseases on Public Health, Greece." <u>Emerging Infectious Diseases</u> **17**(9): 1581-1590.

Glass, G. E., J. E. Cheek, et al. (2000). "Using remotely sensed data to identify areas at risk for hantavirus pulmonary syndrome." <u>Emerging Infectious Diseases</u> **6**(3): 238.

Gostin, L. O. (2005). "World health law: toward a new conception of global health governance for the 21st century." <u>Yale J. Health Pol'y L. & Ethics</u> **5**: 413.

Goutard, F., F. Roger, et al. (2007). "Conceptual framework for avian influenza risk assessment in Africa: The case of Ethiopia." <u>Avian Diseases</u> **51**(1): 504-506.

Goutard, F. L., M. Paul, et al. (2012). "Optimizing early detection of avian influenza H5N1 in backyard and free-range poultry production systems in Thailand." <u>Preventive Veterinary Medicine</u> **105**(3): 223-234.

Grange, Z. L., M. Van Andel, et al. (2014). "Network Analysis of Translocated Takahe Populations to Identify Disease Surveillance Targets." <u>Conservation Biology</u> **28**(2): 518-528.

Grenfell, B., O. Bjørnstad, et al. (2001). "Travelling waves and spatial hierarchies in measles epidemics." <u>nature</u> **414**(6865): 716-723.

Gubler, D. J. (1998). "Resurgent vector-borne diseases as a global health problem." <u>Emerging Infectious Diseases</u> **4**(3): 442.

Gubler, D. J. (2007). "The continuing spread of West Nile virus in the western hemisphere." <u>Clinical Infectious Diseases</u> **45**(8): 1039-1046.

Gubler, D. J. (2012). "The economic burden of dengue." <u>The American journal of Tropical Medicine and Hygiene</u> **86**(5): 743-744.

Gustafson, L., K. Klotins, et al. (2010). "Combining surveillance and expert evidence of viral hemorrhagic septicemia freedom: a decision science approach." <u>Preventive Veterinary Medicine</u> **94**(1-2): 140-153.

Hadorn, D. C., V. Racloz, et al. (2009). "Establishing a cost-effective national surveillance system for Bluetongue using scenario tree modelling." <u>Veterinary Research</u> **40**(6).

Hadorn, D. C., J. Rufenacht, et al. (2002). "Risk-based design of repeated surveys for the documentation of freedom from non-highly contagious diseases." Preventive Veterinary Medicine **56**(3): 179-192.

Halliday, J. E., A. L. Meredith, et al. (2007). "A framework for evaluating animals as sentinels for infectious disease surveillance." Journal of the Royal Society Interface **4**(16): 973-984.

Hamede, R. K., J. Bashford, et al. (2009). "Contact networks in a wild Tasmanian devil (Sarcophilus harrisii) population: using social network analysis to reveal seasonal variability in social behaviour and its implications for transmission of devil facial tumour disease." Ecology letters **12**(11): 1147-1157.

Hay, S., A. Tatem, et al. (2006). "Global environmental data for mapping infectious disease distribution." Advances in Parasitology **62**: 37-77.

Hay, S. I., K. E. Battle, et al. (2013). "Global mapping of infectious disease." Philosophical Transactions of the Royal Society of London B: Biological Sciences **368**(1614): 20120250.

Hay, S. I., D. B. George, et al. (2013). "Big Data Opportunities for Global Infectious Disease Surveillance." Plos Medicine **10**(4).

Hay, S. I., C. A. Guerra, et al. (2009). "A world malaria map: Plasmodium falciparum endemicity in 2007." Plos Medicine **6**(3): 286.

Healy, J. M., W. K. Reisen, et al. (2015). "Comparison of the Efficiency and Cost of West Nile Virus Surveillance Methods in California." Vector-Borne and Zoonotic Diseases **15**(2): 147-155.

Heckathorn, D. D. (1997). "Respondent-driven sampling: a new approach to the study of hidden populations." SOCIAL PROBLEMS-NEW YORK- **44**: 174-199.

Herbreteau, V., G. Salem, et al. (2007). "Thirty years of use and improvement of remote sensing, applied to epidemiology: from early promises to lasting frustration." Health & Place **13**(2): 400-403.

Heymann, D. L. and G. R. Rodier (2001). "Hot spots in a wired world: WHO surveillance of emerging and re-emerging infectious diseases." The Lancet Infectious Diseases **1**(5): 345-353.

Hii, Y. L., J. Rocklov, et al. (2012). "Optimal Lead Time for Dengue Forecast." Plos Neglected Tropical Diseases **6**(10).

Hirzel, A. H., J. Hausser, et al. (2002). "Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data?" Ecology **83**(7): 2027-2036.

Hoinville, L., A. Ronello, et al. (2011). "Animal health surveillance terminology final report from Pre-ICAHS workshop." Health **2011**: 1-26.

Hoinville, L. J., L. Alban, et al. (2013). "Proposed terms and concepts for describing and evaluating animal-health surveillance systems." Preventive Veterinary Medicine **112**(1-2): 1-12.

Huisman, M. and M. A. Van Duijn (2005). "Software for social network analysis." Models and Methods in SocialNetwork Analysis: 270-316.

Isella, L., M. Romano, et al. (2011). "Close encounters in a pediatric ward: measuring face-to-face proximity and mixing patterns with wearable sensors." <u>Plos One</u> **6**(2): e17144.

Janjua, N. Z., D. M. Skowronski, et al. (2012). "Estimates of influenza vaccine effectiveness for 2007–2008 from Canada's sentinel surveillance system: cross-protection against major and minor variants." <u>Journal of Infectious Diseases</u>: jis283.

Jones, K. E., N. G. Patel, et al. (2008). "Global trends in emerging infectious diseases." <u>Nature</u> **451**(7181): 990-993.

Kadohira, M., G. Hill, et al. (2015). "Stakeholder prioritization of zoonoses in Japan with analytic hierarchy process method." <u>Epidemiology and Infection</u> **143**(7): 1477-1485.

Kalluri, S., P. Gilruth, et al. (2007). "Surveillance of arthropod vector-borne infectious diseases using remote sensing techniques: a review." <u>PLoS Pathog</u> **3**(10): 1361-1371.

Karesh, W. B., A. Dobson, et al. (2012). "Ecology of zoonoses: natural and unnatural histories." <u>The lancet</u> **380**(9857): 1936-1945.

Karon, J. M. and C. Wejnert (2012). "Statistical methods for the analysis of time–location sampling data." <u>Journal of Urban Health</u> **89**(3): 565-586.

Kawachi, I. and S. Wamala (2006). <u>Globalization and Health</u>, Oxford University Press.

Keeling, M. J., M. E. Woolhouse, et al. (2001). "Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape." <u>Science</u> **294**(5543): 813-817.

Kigadye, R.-M., A. Klokke, et al. (1993). "Sentinel surveillance for HIV-1 among pregnant women in a developing country: 3 years' experience and comparison with a population serosurvey." <u>Aids</u> **7**(6): 849-856.

Kilpatrick, A. M., A. A. Chmura, et al. (2006). "Predicting the global spread of H5N1 avian influenza." <u>Proceedings of the National Academy of Sciences</u> **103**(51): 19368-19373.

Kitron, U. (2000). "Risk maps: transmission and burden of vector-borne diseases." <u>Parasitology Today</u> **16**(8): 324-325.

Klovdahl, A. S. (1985). "Social networks and the spread of infectious diseases: the AIDS example." <u>Social Science & Medicine</u> **21**(11): 1203-1216.

Knight-Jones, T. J. D., R. Hauser, et al. (2010). "Evaluation of effectiveness and efficiency of wild bird surveillance for avian influenza." <u>Veterinary Research</u> **41**(4).

Komar, N. (2001). "West Nile virus surveillance using sentinel birds." <u>Annals of the New York Academy of Sciences</u> **951**(1): 58-73.

Krause, G. (2008). "Prioritisation of infectious diseases in public health-call for comments."

Krause, J., S. Krause, et al. (2013). "Reality mining of animal social systems." <u>Trends in Ecology & Evolution</u> **28**(9): 541-551.

Kuiken, T., F. Leighton, et al. (2005). "Pathogen surveillance in animals." <u>Science</u> **309**(5741): 1680.

Kulasekera, V. L., L. Kramer, et al. (2001). "West Nile virus infection in mosquitoes, birds, horses, and humans, Staten Island, New York, 2000." <u>Emerging Infectious Diseases</u> **7**(4): 722.

Kung, N. Y., R. S. Morris, et al. (2007). "Risk for infection with highly pathogenic influenza A virus (H5N1) in chickens, Hong Kong, 2002." Emerging Infectious Diseases **13**(3): 412.

L'Herminez, M. and M. T. Mbizvo (1997). "Sentinel surveillance identifies a high HIV prevalence in a rural population in Zimbabwe." Central African Journal of Medicine **43**(4): 103-107.

Langmuir, A. D. (1963). "The surveillance of communicable diseases of national importance." New England journal of medicine **268**(4): 182-192.

Lawson, A., A. Biggeri, et al. (1999). Disease mapping and risk assessment for public health, John Wiley & Sons.

Lenaway, D. D. and A. Ambler (1995). "Evaluation of a school-based influenza surveillance system." Public Health Reports **110**(3): 333.

Levin, S. A., B. Grenfell, et al. (1997). "Mathematical and computational challenges in population biology and ecosystems science." Science **275**(5298): 334-343.

Lewis, H. (2008). One world, one health, one medicine: From the perspective of companion animal practice. Proceedings of the 2008 World Veterinary Congress.

Linthicum, K. J., A. Anyamba, et al. (1999). "Climate and satellite indicators to forecast Rift Valley fever epidemics in Kenya." Science **285**(5426): 397-400.

Lusseau, D., H. Whitehead, et al. (2008). "Incorporating uncertainty into the study of animal social networks." Animal Behaviour **75**(5): 1809-1815.

MacDiarmid, S. (1988). "Future options for brucellosis surveillance in New Zealand beef herds." New Zealand Veterinary Journal **36**(1): 39-42.

Magnani, R., K. Sabin, et al. (2005). "Review of sampling hard-to-reach and hidden populations for HIV surveillance." Aids **19**: S67-S72.

Mak, S., M. Morshed, et al. (2010). "Ecological Niche Modelling of Lyme Disease in British Columbia, Canada." Journal of Medical Entomology **47**(1): 99-105.

Malekinejad, M., L. G. Johnston, et al. (2008). "Using respondent-driven sampling methodology for HIV biological and behavioral surveillance in international settings: a systematic review." AIDS and Behavior **12**(1): 105-130.

Marano, N., P. M. Arguin, et al. (2007). "Impact of globalization and animal trade on infectious disease ecology." Emerging Infectious Diseases **13**(12): 1807.

Marmot, M. and R. Wilkinson (2005). Social determinants of health, Oxford University Press.

Martin, P., A. Cameron, et al. (2007a). "Demonstrating freedom from disease using multiple complex data sources: 2: Case study—Classical swine fever in Denmark." Preventive Veterinary Medicine **79**(2): 98-115.

Martin, P., A. Cameron, et al. (2007b). "Demonstrating freedom from disease using multiple complex data sources: 2: Case study—Classical swine fever in Denmark." Preventive Veterinary Medicine **79**(2): 98-115.

Martin, P., A. Cameron, et al. (2007b). "Demonstrating freedom from disease using multiple complex data sources: 1: A new methodology based on scenario trees." Preventive Veterinary Medicine **79**(2): 71-97.

Martin, V., D. U. Pfeiffer, et al. (2011). "Spatial distribution and risk factors of highly pathogenic avian influenza (HPAI) H5N1 in China." PLoS Pathog **7**(3): e1001308.

Martin, V., X. Zhou, et al. (2011). "Risk-based surveillance for avian influenza control along poultry market chains in South China: The value of social network analysis." Preventive Veterinary Medicine **102**(3): 196-205.

Martinez-Lopez, B., A. M. Perez, et al. (2009). "Social Network Analysis. Review of General Concepts and Use in Preventive Veterinary Medicine." Transboundary and Emerging Diseases **56**(4): 109-120.

McCluskey, B. (2003). "8Chapter." Animal Disease Surveillance and Survey Systems: 119.

McCluskey, B. (2003). "Use of sentinel herds to study the epidemiology of vesicular stomatitis in the state of Colorado." Animal Disease Surveillance and Survey Systems: 119.

McInnes, C. and K. Lee (2006). "Health, security and foreign policy." Review of International Studies **32**(01): 5-23.

McKenzie, J., H. Simpson, et al. (2007). "Development of methodology to prioritise wildlife pathogens for surveillance." Preventive Veterinary Medicine **81**(1-3): 194-210.

McMichael, A. J. (2004). "Environmental and social influences on emerging infectious diseases: past, present and future." Philosophical Transactions of the Royal Society B: Biological Sciences **359**(1447): 1049-1058.

Minh, P., B. Schauer, M. Stevenson, G. Jones, R. Morris and A. Noble (2009). "Association between human cases and poultry outbreaks of highly pathogenic avian influenza in Vietnam from 2003 to 2007: a nationwide study." Transboundary and emerging diseases 56(8): 311-320.

Morris, M., J. Zavisca, et al. (1995). "Social and sexual networks: their role in the spread of HIV/AIDS among young gay men." AIDS Education and Prevention.

Mostashari, F., M. Kulldorff, et al. (2003). "Dead bird clusters as an early warning system for West Nile virus activity." Emerging Infectious Diseases **9**(6): 641.

Murphy, T., H. Wahlström, et al. (2012). "Freedom from Echinococcus multilocularis: an Irish perspective." Veterinary Parasitology **190**(1): 196-203.

Najjar, J. K. (2014). "Planning for Big Data: A CIO's Handbook to the Changing Data Landscape."

Neslo, R. and R. Cooke (2011). "Modelling and validating stakeholder preferences with probabilistic inversion." Applied Stochastic Models in Business and Industry **27**(2): 115-130.

Newman, M. E. (2003). "Properties of highly clustered networks." Physical Review E **68**(2): 026121.

Ng, V. and J. M. Sargeant (2013). "A quantitative approach to the prioritization of zoonotic diseases in North America: a health professionals' perspective." Plos One **8**(8): e72172.

Noor, A. M., D. K. Kinyoki, et al. (2014). "The changing risk of Plasmodium falciparum malaria infection in Africa: 2000-10: a spatial and temporal analysis of transmission intensity." Lancet **383**(9930): 1739-1747.

Noremark, M., N. Hakansson, et al. (2011). "Network analysis of cattle and pig movements in Sweden: Measures relevant for disease control and risk-based surveillance." Preventive Veterinary Medicine **99**(2-4): 78-90.

Organization, W. H. (2003). "The Dubrovnik pledge on surveillance and prioritization of infectious diseases: report on a WHO meeting, Bucharest, Romania 21-23 November 2002."

Ortiz-Pelaez, A. and D. U. Pfeiffer (2008). "Use of data mining techniques to investigate disease risk classification as a proxy for compromised biosecurity of cattle herds in wales." BMC Veterinary Research **4**(24): (04 July 2008)-(2004 July 2008).

Ortiz-Pelaez, A., D. U. Pfeiffer, et al. (2006). "Use of social network analysis to characterize the pattern of animal movements in the initial phases of the 2001 foot and mouth disease (FMD) epidemic in the UK." Preventive Veterinary Medicine **76**(1-2): 40-55.

Otte, E. and R. Rousseau (2002). "Social network analysis: a powerful strategy, also for the information sciences." Journal of Information Science **28**(6): 441-453.

Paige, J. C., M. H. Chaudry, et al. (1999). "Federal surveillance of veterinary drugs and chemical residues (with recent data)." The Veterinary Clinics of North America. Food Animal Practice **15**(1): 45-61, viii.

Patz, J. A., P. Daszak, et al. (2004). "Unhealthy landscapes: policy recommendations on land use change and infectious disease emergence." Environmental Health Perspectives **112**(10): 1092.

Paul, M., L. Held, et al. (2008). "Multivariate modelling of infectious disease surveillance data." Statistics in Medicine **27**(29): 6250-6267.

Pedley, S. and K. Pond (2003). Emerging Issues in Water and infectious Disease, World Health Organization.

Phillips, S. J., R. P. Anderson, et al. (2006). "Maximum entropy modelling of species geographic distributions." Ecological Modelling **190**(3): 231-259.

Piot, P., J.-J. Muyembe, et al. (2014). "Ebola in west Africa: from disease outbreak to humanitarian crisis." The Lancet Infectious Diseases **14**(11): 1034-1035.

Pozio, E., L. Alban, et al. (2010). Development of harmonised schemes for the monitoring and reporting of Trichinella in animals and foodstuffs in the European Union. SCIENTIFIC REPORT submitted to EFSA.

Prah, P., A. J. Copas, et al. (2013). "Consistency in reporting sensitive sexual behaviours in Britain: change in reporting bias in the second and third National Surveys of Sexual Attitudes and Lifestyles (Natsal-2 and Natsal-3)." Sexually Transmitted Infections.

Prattley, D. J., R. S. Morris, et al. (2007). "Application of portfolio theory to risk-based allocation of surveillance resources in animal populations." Preventive Veterinary Medicine **81**(1-3): 56-69.

Presi, P., K. D. C. Staerk, et al. (2008). "Efficiency of risk-based vs. random sampling for the monitoring of tetracycline residues in slaughtered calves in Switzerland." Food Additives and Contaminants **25**(5): 566-573.

Presi, P., K. D. Stärk, et al. (2009). "Risk scoring for setting priorities in a monitoring of antimicrobial resistance in meat and meat products." International Journal of Food Microbiology **130**(2): 94-100.

Rabinowitz, P. M., Z. Gordon, et al. (2005). "Animals as sentinels of human environmental health hazards: an evidence-based analysis." EcoHealth **2**(1): 26-37.

Racloz, V., C. Griot, et al. (2006). "Sentinel surveillance systems with special focus on vector-borne diseases." Animal Health Research Reviews **7**(1-2): 71-79.

Racloz, V., R. Straver, et al. (2006). "Establishment of an early warning system against Bluetongue virus in Switzerland." Schweizer Archiv Fur Tierheilkunde **148**(11): 593-598.

Racloz, V., G. Venter, et al. (2008). "Estimating the temporal and spatial risk of bluetongue related to the incursion of infected vectors into Switzerland." BMC Veterinary Research **4**.

Raska, K. (1966). "National and international surveillance of communicable diseases."

Rautureau, S., B. Dufour, et al. (2011). "Targeted surveillance of cattle trade using social network analysis tools." Epidemiologie et Sante Animale(59/60): 58-60.

Raymond, H. F., T. Ick, et al. (2007). Resource guide: time location sampling (TLS), San Francisco, CA: San Francisco Department of Public Health HIV Epidemiology Section, Behavioral Surveillance Unit.

Reintjes, R. and L. Wiessing (2007). "2nd-generation HIV surveillance and injecting drug use: uncovering the epidemiological ice-berg." International Journal of Public Health **52**(3): 166-172.

Reist, M., T. Jemmi, et al. (2012). "Policy-driven development of cost-effective, risk-based surveillance strategies." Preventive Veterinary Medicine **105**(3): 176-184.

Ribeiro-Lima, J., E. A. Enns, et al. (2015). "From network analysis to risk analysis-An approach to risk-based surveillance for bovine tuberculosis in Minnesota, US." Preventive Veterinary Medicine **118**(4): 328-340.

Ribeiro, J., F. Seulu, et al. (1996). "Temporal and spatial distribution of anopheline mosquitos in an Ethiopian village: implications for malaria control strategies." Bulletin of the World Health Organization **74**(3): 299.

Ricci, A., K. Capello, et al. (2013). "Raw milk-associated foodborne infections: A scoring system for the risk-based categorisation of raw dairy farms." Research in Veterinary Science **95**(1): 69-75.

Richard, J.-L., B. Vidondo, et al. (2008). "A 5-year comparison of performance of sentinel and mandatory notification surveillance systems for measles in Switzerland." European Journal of Epidemiology **23**(1): 55-65.

Richmond, C., S. J. Elliott, et al. (2005). "The political ecology of health: perceptions of environment, economy, health and well-being among 'Namgis First Nation." Health & Place **11**(4): 349-365.

Roberts, R. S. and I. M. Foppa (2006). "Prediction of equine risk of West Nile Virus infection based on dead bird surveillance." Vector-Borne and Zoonotic Diseases **6**(1): 1-6.

Robertson, C., T. A. Nelson, et al. (2010). "Review of methods for space–time disease surveillance." Spatial and Spatio-temporal Epidemiology **1**(2): 105-116.

Robinson, A., M. A. Burgman, et al. (2011). "Allocating surveillance resources to reduce ecological invasions: maximizing detections and information about the threat." Ecological Applications : A Publication of the Ecological Society of America **21**(4): 1410-1417.

Robinson, S. and R. Christley (2007). "Exploring the role of auction markets in cattle movements within Great Britain." Preventive Veterinary Medicine **81**(1): 21-37.

Rodríguez-Prieto, V., M. Vicente-Rubiano, et al. (2014). "Systematic review of surveillance systems and methods for early detection of exotic, new and re-emerging diseases in animal populations." Epidemiology and Infection: 1-25.

Rogers, D. (2006). "Models for vectors and vector-borne diseases." Advances in Parasitology **62**: 1-35.

Rogers, D. J. and S. E. Randolph (2003). "Studying the global distribution of infectious diseases using GIS and RS." Nature Reviews Microbiology **1**(3): 231-237.

Rogers, D. J., S. E. Randolph, et al. (2002). "Satellite imagery in the study and forecast of malaria." Nature **415**(6872): 710-715.

Rogers, D. J., C. Tucker, et al. (2002). "Predicting the distribution of West Nile fever in North America using satellite sensor data." Photogrammetric Engineering and Remote Sensing **68**(2): 112-136.

Rushdy, A. and M. O'Mahony (1998). "PHLS overview of communicable diseases 1997: results of a priority setting exercise." Communicable Disease Report. CDR weekly **8**: S1-12.

Ryan, C. P. (2008). "Zoonoses likely to be used in bioterrorism." Public health reports **123**(3): 276.

Sachs, J. and P. Malaney (2002). "The economic and social burden of malaria." Nature **415**(6872): 680-685.

Salathé, M., M. Kazandjieva, et al. (2010). "A high-resolution human contact network for infectious disease transmission." Proceedings of the National Academy of Sciences **107**(51): 22020-22025.

Samat, N. A. and S. H. M. I. Ma'arof (2014). "Disease Mapping based on Stochastic SIR-SI Model for Dengue and Chikungunya in Malaysia." International Conference on Quantitative Sciences and Its Applications (Icoqsia 2014) **1635**: 227-234.

53

Sanchez-Matamoros, A., B. Martinez-Lopez, et al. (2013). "Social network analysis of Equidae movements and its application to risk-based surveillance and to control of spread of potential Equidae diseases." Transboundary and Emerging Diseases **60**(5): 448-459.

Schwarcz, S. K., J. M. Zenilman, et al. (1990). "National surveillance of antimicrobial resistance in Neisseria gonorrhoeae." JAMA **264**(11): 1413-1417.

Schwermer, H., I. Reding, et al. (2009). "Risk-based sample size calculation for consecutive surveys to document freedom from animal diseases." Preventive Veterinary Medicine **92**(4): 366-372.

Sergeant, E., A. Cameron, et al. (2009). "Scenario trees to quantify freedom from disease." Web-based software accessed at http://freedom. ausvet. com. au.

Sergeant ESG; Cameron, A. H., JM; Martin, PAJ. (1990). "Scenario trees to quantify freedom from disease." 2015.

Shannon, G. W. (1981). "Disease mapping and early theories of yellow fever." The Professional Geographer **33**(2): 221-227.

Shirley, M. D. and S. P. Rushton (2005). "The impacts of network topology on disease spread." Ecological Complexity **2**(3): 287-299.

Skowronski, D., N. Janjua, et al. (2013). "Interim estimates of influenza vaccine effectiveness in 2012/13 from Canada's sentinel surveillance network, January 2013." Euro Surveill **18**(5): 20394.

Skowronski, D. M., G. De Serres, et al. (2009). "Component-specific effectiveness of trivalent influenza vaccine as monitored through a sentinel surveillance network in Canada, 2006–2007." Journal of Infectious Diseases **199**(2): 168-179.

Slingenbergh, J., M. Gilbert, et al. (2004). "Ecological sources of zoonotic diseases." Revue Scientifique Et Technique-Office International Des Epizooties **23**(2): 467-484.

Smieszek, T., E. Burri, et al. (2012). "Collecting close-contact social mixing data with contact diaries: reporting errors and biases." Epidemiology and Infection **140**(04): 744-752.

Smieszek, T., L. Fiebig, et al. (2009). "Models of epidemics: when contact repetition and clustering should be included." Theoretical Biology and Medical Modelling **6**(1): 11.

Smieszek, T. and M. Salathé (2013). "A low-cost method to assess the epidemiological importance of individuals in controlling infectious disease outbreaks." BMC medicine **11**(1): 35.

Snow, L. C., S. E. Newson, et al. (2007). "Risk-based surveillance for H5N1 avian influenza virus in wild birds in Great Britain." The Veterinary Record **161**(23): 775-781.

Soh, S. E., A. R. Cook, et al. (2012). "Teacher led school-based surveillance can allow accurate tracking of emerging infectious diseases - evidence from serial cross-sectional surveys of febrile respiratory illness during the H1N1 2009 influenza pandemic in Singapore." BMC Infectious Diseases **12**.

Spiller, M., C. Cameron, et al. (2012). "RDSAT 7.1 user manual." Cornell University.

Stark, K. D. C., G. Regula, et al. (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): (28 February 2006)-(2028 February 2006).

States, U. (2000). "National intelligence estimate: the global infectious disease threat and its implications for the United States." Environmental Change and Security Project report(6): 33.

Stebler, N., G. Schuepbach-Regula, et al. (2015). "Use of a modified Delphi panel to identify and weight criteria for prioritization of zoonotic diseases in Switzerland." Preventive Veterinary Medicine **121**(1-2): 165-169.

Stockwell, D. (1999). "The GARP modelling system: problems and solutions to automated spatial prediction." International Journal of Geographical Information Science **13**(2): 143-158.

Suarez, D. L. (2005). "Overview of avian influenza DIVA test strategies." Biologicals **33**(4): 221-226.

Sumption, K., M. Rweyemamu, et al. (2008). "Incidence and Distribution of Foot-and-Mouth Disease in Asia, Africa and South America; Combining Expert Opinion, Official Disease Information and Livestock Populations to Assist Risk Assessment." Transboundary and Emerging Diseases **55**(1): 5-13.

Sutherst, R., G. Maywald, et al. (2007). "CLIMEX version 3: user's guide." Hearne Scientific Software Pty Ltd.

Tavernier, P., J. Dewulf, et al. (2011). "Wildtool, a flexible, first-line risk assessment system for wildlife-borne pathogens." European Journal of Wildlife Research **57**(5): 1065-1075.

Tavornpanich, S., I. A. Gardner, et al. (2006). "Evaluation of cost-effectiveness of targeted sampling methods for detection of Mycobacterium aviumsubsp paratuberculosis infection in dairy herds." American Journal of Veterinary Research **67**(5): 821-828.

Taylor, L. H., S. M. Latham, et al. (2001). "Risk factors for human disease emergence." Philosophical Transactions of the Royal Society of London B: Biological Sciences **356**(1411): 983-989.

Thacker, S. B. and R. L. Berkelman (1988). "Public health surveillance in the United States." Epidemiologic Reviews **10**: 164-190.

Tsai, P., K. A. Scott, et al. (2009). Sustaining Global Surveillance and Response to Emerging Zoonotic Diseases, National Academies Press.

UNDG (2015). Socio-Economic Impact of Ebola Virus Disease in West African Countries - A call for national and regional containment, recovery and prevention. New York, USA, United Nations Development Programme Regional Bureau for Africa**:** 93.

Van der Schalie, W. H., H. S. Gardner Jr, et al. (1999). "Animals as sentinels of human health hazards of environmental chemicals." Environmental Health Perspectives **107**(4): 309.

Van Kerkhove, M. D., S. Vong, et al. (2009). "Poultry movement networks in Cambodia: implications for surveillance and control of highly pathogenic avian influenza (HPAI/H5N1)." Vaccine **27**(45): 6345-6352.

Volkova, V., P. Bessell, et al. (2011). "Paper: Evaluation of risks of foot-and-mouth disease in Scotland to assist with decision making during the 2007 outbreak in the UK." The Veterinary Record **169**(5): 124.

Volz, E. and D. D. Heckathorn (2008). "Probability based estimation theory for respondent driven sampling." Journal of Official Statistics **24**(1): 79.

Wagner, M. M., A. W. Moore, et al. (2011). Handbook of biosurveillance, Academic Press.

Wahlström, H., J. Frössling, et al. (2010). "Demonstrating freedom from Mycobacterium bovis infection in Swedish farmed deer using non-survey data sources." Preventive Veterinary Medicine **94**(1): 108-118.

Walsh, D. P. and M. W. Miller (2010). "A Weighted Surveillance Approach for Detecting Chronic Wasting Disease Foci." Journal of Wildlife Diseases **46**(1): 118-135.

Wasserman, S. and K. Faust (1994). Social Network Analysis: Methods and Applications, Cambridge university press.

Wattana, W., F. van Griensven, et al. (2007). "Respondent-driven sampling to assess characteristics and estimate the number of injection drug users in Bangkok, Thailand." Drug and Alcohol Dependence **90**(2): 228-233.

Wei, C., W. McFarland, et al. (2012). "Reaching black men who have sex with men: a comparison between respondent-driven sampling and time-location sampling." Sexually Transmitted Infections: sextrans-2012-050619.

Weinberg, J., O. Grimaud, et al. (1999). "Establishing priorities for European collaboration in communicable disease surveillance." The European Journal of Public Health **9**(3): 236-240.

Weinberg, M., S. Waterman, et al. (2003). "The US-Mexico border infectious disease surveillance project: Establishing binational border surveillance." Emerging Infectious Diseases **9**(1): 97.

Welby, S., E. Meroc, et al. (2013). "Bluetongue surveillance system in Belgium: A stochastic evaluation of its risk-based approach effectiveness." Preventive Veterinary Medicine **112**(1-2): 48-57.

Whitehead, H. (2008). Analyzing animal societies: quantitative methods for vertebrate social analysis, University of Chicago Press.

WHO. (2004). "WHO guidelines for global surveillance of influenza A/H5." Retrieved 8/14, 2015, from http://www.who.int/influenza/surveillance_monitoring/global_surveillance_h5_guidelines_06_02_2004/en/

WHO. (2008). "International Health Regulations (2005) -- Second Edition." Retrieved 7/6, 2015, from http://whqlibdoc.who.int/publications/2008/9789241580410_eng.pdf?ua=1

Wilcox, B. A. and D. J. Gubler (2005). "Disease ecology and the global emergence of zoonotic pathogens." Environmental Health and Preventive Medicine **10**(5): 263-272.

Willeberg, P., L. R. Nielsen, et al. (2012). "Designing and evaluating risk-based surveillance systems: Potential unwarranted effects of applying adjusted risk estimates." Preventive Veterinary Medicine **105**(3): 185-194.

Winkelstein, W., D. M. Lyman, et al. (1987). "Sexual practices and risk of infection by the human immunodeficiency virus." JAMA **257**(3): 321-325.

Wiratsudakul, A., M. C. Paul, et al. (2014). "Modelling the dynamics of backyard chicken flows in traditional trade networks in Thailand: implications for surveillance and control of avian influenza." Tropical Animal Health and Production **46**(5): 845-853.

Witt, C. J., A. L. Richards, et al. (2011). "The AFHSC-Division of GEIS Operations Predictive Surveillance Program: a multidisciplinary approach for the early detection and response to disease outbreaks." BMC Public Health **11**.

Wood, B. L., L. R. Beck, et al. (1991). "Spectral and spatial characterization of rice field mosquito habitat." International Journal of Remote Sensing **12**(3): 621-626.

Woolhouse, M. (2011). "How to make predictions about future infectious disease risks." Philosophical Transactions of the Royal Society B: Biological Sciences **366**(1573): 2045-2054.

Yang, G.-J., P. Vounatsou, et al. (2005). "A review of geographic information system and remote sensing with applications to the epidemiology and control of schistosomiasis in China." Acta Tropica **96**(2): 117-129.

Zinsstag, J., E. Schelling, D. Waltner-Toews and M. Tanner (2011). "From "one medicine" to "one health" and systemic approaches to health and well-being." Preventive Veterinary Medicine 101(3): 148-156.

3. From Science to Application: Development of a User-friendly Program to Facilitate Risk-based Approaches for Disease Surveillance - HandiResponse

3.1. Abstract

Background: The increasing disease burden from emerging infectious diseases and their impacts calls for more effective surveillance approaches. Risk-based surveillance and integration of multiple forms of surveillance are prominent in current suggestions on how to enhance disease surveillance strategies. However, quite a number of constraints prevent countries, especially resource poor ones, from adopting them.

Method: Based upon the results from literature review, the architecture of a disease surveillance planning tool was designed. All modules are developed in a way that outputs from the modules in the upstream are used as the inputs for the modules in the downstream.

Results: Human and Animal Disease Response Program (HandiReponse) has been designed to help assess and improve disease surveillance programs. The program includes four modules: (i) risk map development – HandiMap; (ii) surveillance portfolio development – HandiSurv; (iii) economic impact assessment – HandiEcon and (iv) surveillance optimization – OptiSurv. The conceptual framework, objectives and structure of the program are described, and examples given of its application.

Conclusion: HandiResponse, combining risk mapping, infectious disease modelling, surveillance program planning and optimization in one package, which can potentially facilitate effective communication between technical staff and policy makers on disease surveillance and improve technical and economic efficiency of the programs for disease surveillance, in particular, those in developing world. Its performance await to be tested and user friendliness to be improved.

3.2. Introduction

Risk-based surveillance has gained in popularity in recent years, and represents an important development for modern disease surveillance (Qing et al., 2003; Stark et al., 2006). Risk-based surveillance programs have been promoted as potentially outperforming the traditional non-risk-based ones in terms of improved system efficacy by reduction of sample size for demonstration of disease freedom, quicker and more accurate detection of disease presence, and improved probability of detection of rare diseases (Hadorn, Rufenacht et al., 2002; Schwermer, Reding, Hadorn et al., 2009; Willeberg et al., 2012). They could also achieve higher benefit-cost ratio or cost saving (Alba et al., 2010; F. F. Calvo-Artavia et al., 2013; Catherine G. Geanuracos et al., 2007; Presi et al., 2008; Reist et al., 2012; Tavornpanich et al., 2006) and detect outbreaks of infectious diseases earlier (Kahn, 2006; Kuiken et al., 2005; Kulasekera et al., 2001; Racloz, Venter et al.,2008). However, planning and design of a risk-based approach for surveillance can be challenging for a number of reasons. The approaches adopted need to overcome policy constraints (Cameron, 2012; Martin, Cameron et al., 2007a; Qing, Saijo et al., 2003; Stark, Regula et al., 2006; Tsai, Scott et al., 2009; Weinberg, Waterman et al., 2003), require access to intensive and up to date data (Alba, Casal et al., 2010; Carroll, Au et al., 2014; Christensen, Stryhn et al., 2014; Kalluri, Gilruth et al., 2007; Ortiz-Pelaez, Pfeiffer et al., 2006) and need adequate epidemiological and research capacity (Kalluri et al., 2007). Implementing the approach may also be expensive (Smieszek & Salathé, 2013). These challenges are more prominent in resource-constrained countries. Hence, most of the risk-based surveillance activities have so far been implemented in developed economies, whereas they could provide greater benefit if applied in resource-constrained countries. However these countries do not have the same level of expertise to apply risk-based methods.

A generic management tool has been developed in order to help assess and improve national disease surveillance programs and the utilization of a risk-based approach - in particular, developing and under-developed countries. The system is called Human and Animal Disease Response System (HandiRepsonse). In the following sections, the conceptual framework, objectives and structure of the program will be described.

3.3. Materials and methods

3.3.1. Program description

3.3.1.1. Conceptual framework

Risk of disease can be defined as the interaction between the probability of disease occurrence and the severity of its consequences (North, 1995). A risk factor is an environmental, behavioral or biological factor that directly increases the probability of a disease occurring (Beck, 1998). Rarely does an emerging infectious disease occur evenly in either spatial or temporal terms, or across population

groups (Stark, Regula et al. 2006). Identification of location, time/season and population at most risk could possibly enable effective use of resources for disease prevention and mitigation. With the advances having taken place in the fields of computer science, remote sensing technologies, geographic information systems, geographic positioning systems, and decision making sciences, the effects of different risk factors can be better gauged and visualized. Such information could help predict infectious disease with sufficient lead time to allow an effective government response to prevent an incursion, or to deal with it effectively after it is introduced (Woolhouse, 2011).

The HandiResponse program employs a sequence of components to develop a "risk landscape" for potential occurrence of a disease of concern, then models the spatial and temporal spread of this disease on the risk landscape, and then uses this model to test the effect of potential surveillance measures. Finally it conducts an optimization procedure to choose a mix of surveillance methods which is expected to be both epidemiologically and economically efficient in detecting an incursion of the disease. Later the program is intended to broaden the range of situations in which it can be used, but an incursion provides a suitable initial case study to test the concept.

HandiResponse uses HandiSpread, a spatial and stochastic simulation modelling program for modelling zoonotic diseases on a user-defined risk landscape, rather than a uniform surface. It was developed from the earlier InterSpread Plus [2] by EpiSoft Company, by adding the capacity to transmit disease to the human population as well as spread within domestic animal and wildlife populations, to create an epidemiological model for a particular disease under study.

Risk assessment and mapping jointly provide the critical first step of producing a risk landscape (map). The disease is then modelled on this risk landscape, and various surveillance methods are tested for their capacity to detect the disease. These results are then used in a further analysis to assess the effect of various potential surveillance approaches on speed of disease detection and cost-effectiveness of a surveillance portfolio consisting of one or more surveillance methods. The reduction in potential consequences measured in economic terms (including less effect on human health), together with the cost for each surveillance approach, are jointly used to construct an optimal surveillance portfolio.

As the name suggests, HandiResponse has been designed with the capacity for managing surveillance program for infectious zoonotic diseases affecting both animals and people.

3.3.1.2. Objectives

The specific objectives of HandiResponse are (i) to visualize the disease risk landscape and identify hotspots where the infectious disease under study is likely to spread most rapidly and hence could be

---

[2]HandiSpread, Massey University EpiCentre, Palmerston North, New Zealand. Available at http://www.interspreadplus.com, accessed on August 14, 2014.

detected by surveillance; (ii) to evaluate economic benefit and costs of each surveillance activity, and their combined effects as components of a portfolio; (iii) to define optimal use of resource for surveillance through selection of an optimal surveillance portfolio (a set of surveillance components and sub-components) across predefined risk categories such as geographic area, species, sectors, types of stakeholders, etc. In this chapter the design concepts of HandiResponse will be described, then in the next Chapter the implementation of the concepts will be demonstrated.

3.3.1.3. Structure

The program comprises four modules (Figure 3-1): (i) risk landscape development – HandiMap; (ii) surveillance method planning – HandiSurv; (iii) economic impact assessment – HandiEcon and (iv) surveillance optimization – OptiSurv. The lines with arrow in the figure indicate the direction of information flow and linkage between different modules. Besides, HandiResponse has to use the disease model HandiSpread to conduct effect evaluation of different surveillance options, which is a critical step for surveillance optimization.
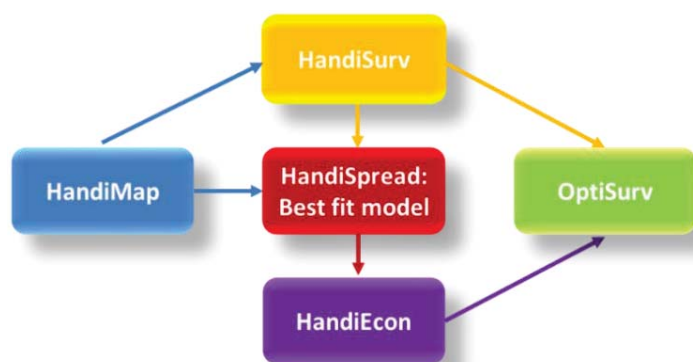
Figure 3-1. Structure of HandiResponse and linkage between different modules

HandiMap Module

This module provides an assessment and mapping tool for estimating and presenting the spatial variation in risk of spread of the disease under study across the defined geographical area by geographic unit. The module will allow a participatory process for identifying key risk factors (and where necessary their environmental or other proxies) and presenting them.

The outputs of the module are (i) risk maps called "risk landscapes" that display the "height" of the risk at each location, and can be single risk factor maps or a combined risk map. These maps can be either in kernel smoothed format or raster format; (ii) a risk score file which transfers the "risk height" information to HandiSpread, for use in adjusting the susceptibility of particular locations to occurrence of the disease, and hence allows HandiSpread to take account of risk level in representing transmission of the disease. A process has been designed for guiding the users through the disease risk mapping process in a stepwise approach (Table 3-1).

Table 3-1. Essential Steps for Risk Mapping by Using HandiMap

Step	Approach	Output
• Screening out areas unsuitable for spread	Areas in which the disease cannot occur are excluded, such as large water bodies, high mountains etc.	Potential suitable locations remain available for mapping
• Defining and selection of risk factors/proxy variables	Defining the epitype (described in the discussion section) of the disease under analysis Selection of risk factors/proxies is informed by evidence from scientific literature, national research, and where appropriate experts' opinions	List of risk factors/proxies
• Establishment of database for analysis	Selection of global spatial data sets already stored within the program, and if desired, import additional data supplied by the user, such as national spatially structured demographic data on animals and people	Spatial dataset for each risk factor/proxy variable
• Risk factor scoring	Standardizing risk factor scores so that all factors are measured on matching scale ranges, so that they can be added together	Standardized score assigned to each spatial unit of observation for each risk factor/proxy
• Prioritization of risk factors	Establishing relative importance of each risk factor by Analytic Hierarchy Process	Weighted standardized score assigned to each spatial unit for each risk factor/proxy
• Summation of all risk scores	Weighted linear combination of the factors using weights from Analytic Hierarchy Process	An overall weighted, standardized and combined risk score assigned to each spatial unit of observation
• Producing risk map	Displaying risk cartography with numerical risk values attached to spatial units	Raster risk maps showing risk of units of observation, or kernel smoothed risk maps
• Calculating risk score for epidemiological spatial units	Defining spatially appropriate epidemiological units for disease and surveillance modelling, such as herds, villages or districts, and calculating the risk value of each of these units	A file containing risk scores for each epidemiological unit.

HandiSurv Module

The module is designed for planning surveillance components and sub-components defined by a program user for a disease and estimating their respective effects on disease detection.

This requires appropriate terminology. For this PhD project, we have defined and used the following technical terms related to disease surveillance.

A surveillance portfolio is a set of disease detection activities carried out in a coordinated fashion. It consists of one or usually multiple **surveillance components**, which are specific surveillance activities defined with regard to method of disease detection, target population and spatial coverage. A surveillance component applies one option selected from multiple **surveillance sub-components**, which usually have at least two dimensions – **surveillance intensity**, which is the level of effort put into the investigation process, and **detection sensitivity**, which is the sensitivity of the diagnostic procedure used in the investigation process. **Compliance** is the third dimension where appropriate. In the example used in Chapters 7 and 8, there are three levels each of surveillance intensity and detection sensitivity, and hence nine combinations of sub-components under each surveillance component. Because of the nature of the surveillance activity there is no compliance sub-component in this case. One of the nine combinations is chosen to be used in the surveillance activity. It would be possible to define sub-components differently if required for a particular situation. An algorithm for surveillance planning and some key terminologies used to describe the composition of a surveillance program are all summarized in Figure 3-2. By the end of the process, subject for surveillance (human, animal or vector, etc.), event for surveillance (syndrome, disease, evidence of infection, etc.), location, method and associated intensity, detection sensitivity and compliance will be defined. Besides, the technical specifications for each subcomponent which needs to be defined are documented in Table 3-2.

The above mentioned terms will be used for defining a surveillance program for a given disease. They need to be parameterized within HandiSurv and input into HandiSpread to estimate their respective effects on disease detection. The effect on disease detection of each individual surveillance subcomponent is estimated separately.

Program
- A plan of action for disease surveillance aims to achieve clear business objectives, with details on what work is to be done, by whom, when, and what means or resources will be used

Event
- Incident
- Syndrome
- Infection
- Disease

Portfolio
- A group of surveillance components and subcomponents

Component
- Species, subgroup
- Property: farm, market, abattoir, clinic, hospital, pharmacy
- Producer, wholesale, retail, restaurant
- Geographic area or season

Subcomponent
- reporting: passive, active
- Syndromic surveillance: passive, active
- Serological surveillance: random sampling, sentinel, risk based
- Pathogen surveillance: random, sampling, sentinel, risk based

Figure 3-2. Algorithm for constructing a surveillance portfolio

Table 3-2. Key Parameters Used for Estimating Performance of Surveillance Subcomponent

Parameter	Explanation	Category
Selection probability	Probability that a surveillance unit will be eligible for being selected under a surveillance subcomponent	Intensity
Sensitivity	the probability that an event will be correctly detected by a surveillance subcomponent	Sensitivity
Visit delay	Probability of a surveillance unit to be visited as a function of the number of days after it has been put on surveillance list	Intensity
Visit frequency	The number of time periods (e.g. days) between visits	Intensity
Visit duration	The number of time periods (e.g. days) a surveillance unit is on surveillance	Intensity
Delay to detection	The number of time periods (e.g. days) between a visit to a surveillance unit to when it is detected	Compliance
Probability of reporting	The probability that a surveillance event will be reported	Compliance
Probability of detection	Probability of a surveillance unit to be detected per unit time (eg. 1 day)	Sensitivity

These technical specifications will be the inputs used in HandiSpread® for estimating impacts of any given surveillance component introduced in the disease model. The default number set for separately seeded replicates of the simulation model is 99. The number of 99 is selected because more than 20 years' experience in using InterSpread Plus suggests such a number of iterations reliably gives results and further increase in iterations provides few marginal benefit in improving modelling results and its associated confidence intervals. Besides, and the number of 99 is chosen so that there is a single median replicate for any variable of interest. The key statistics which need to be reported out of the simulation include a list of the infected properties, the number of animals and people associated with each of the infected properties, and the day on which the disease is detected in each iteration. Each model is run for 365 days or any predefined time period, and the proportion of iterations in which the disease is not detected by the chosen surveillance activities is reported. Further information on HandiSpread and the process of using it for disease modelling is provided in Chapter 5, where the software is used for the first time, and in Chapter 7 where it is used for a second part of the research.

HandiEcon Module

HandiEcon is designed for benefit and cost estimation for any given interventions such as surveillance. The outputs of the module will be cost for each surveillance component and selected subcomponent, gross benefit and net benefit of applying a given surveillance program. Gross benefit for any given surveillance component is defined as the difference between the economic effect of an undetected disease outbreak over 365 days and the economic effect of the outbreak up to the point of detection by the surveillance activity. The net benefit is defined as gross benefit minus the cost for any given surveillance program by the time of detection. A budgeting model has been developed to estimate productivity impact because of disease. The consequences of human disease are estimated in two steps: disease burden is estimated by using WHO Disability Adjusted Life Year (DALY) template; the DALY loss is then converted into economic loss by using the human capital approach. In addition, direct costs of medical care and related expenses for affected people are calculated. All estimation templates are in MS-Excel format.

OptiSurv Module

OptiSurv is an optimization program which can be used to search for top-ranked surveillance portfolios using predefined criteria, be it high technical efficiency, high economic efficiency, high speed of detection of the disease or other criteria. This module helps to answer key questions such as where (geographic location, stage in the market chain, etc.) and what (species, surveillance method, etc.) to spend surveillance resources on. It is in MS-Excel format.

The OptiSurv workbook has three data sheets (*Strategies, FarmCosts and DayDetected*) that contain the data about the simulated epidemic and surveillance, a *Summary* sheet that records your progress in

setting up an optimization, the *Compare* sheet that let you compare five portfolios, and the *Best* sheet that shows the results of the optimization.

Important set of inputs needed for this module include: (i) a list of possible surveillance components and subcomponents that are under consideration to be implemented. (ii) the metrics generated from each of the 99 simulation replicates such as number of days that the particular surveillance component runs until it detects the disease, the number of herds or other population units infected up to the time of disease detection, the cost of surveillance as well as gross and net benefit of detecting the disease compared with an undetected equivalent outbreak where no surveillance was conducted.

OptiSurv calculates the necessary statistics for each portfolio and uses a search algorithm to identify the top ranked surveillance portfolios according to the particular decision rule that has been set. OptiSurv evaluates each possible portfolio according to the decision rule, and excludes clearly suboptimal clusters of portfolios as it narrows its search to produce a ranked list of the portfolios which best satisfy the decision rule.

3.3.2. Program flow

As illustrated in Figure 3-1, the program starts with HandiMap, which produces risk maps and generates a risk file with a score for each spatial unit of epidemiological importance. The information on risk scores is then fed into HandiSpread to develop a temporal and spatial model representing the spread of the disease of concern following an incursion - in the first instance without any detection methods in place, so the disease remains unidentified. To represent the natural variation in spread pattern of the disease from the same incursion point, 99 replicate epidemics are simulated, with different random number seeds, so that biological variability within the same parameter settings is represented. The structure of HandiSpread allows the same 99 replicate epidemics to be run as many times as needed, with exactly the same disease behaviour, but with separately seeded surveillance strategies applied to the disease so that surveillance strategies can be compared on equal terms. The model is used to assess the detection effectiveness of each surveillance component/subcomponent combination. The metrics generated from the simulation are transferred into HandiEcon for estimation of the economic effect of each surveillance component/sub-component combination. Finally, the statistics produced by HandiSpread and HandiEcon are used by OptiSurv to identify optimal surveillance portfolios from within all the combinations available.

3.4. Implementation

In Chapter 3 the software tools are illustrated, showing the way in which they have been used to implement the approach described in this Chapter. Then in Chapters 4 to 8 practical applications of

the various tools to stages of the HandiResponse surveillance design process are illustrated with available data sets.

3.5. Discussion

3.5.1. Program design

The program has been designed to provide for a consensus-building process that allows participatory decision making about surveillance for diseases of national concern. The investigation team can, for instance, debate the selection of risk factors and the weight to be given to each individual risk factor/layer across different risk factors. The products of the program include a risk map, and a selection of optimal surveillance portfolios linked to alternative decision rules which are suited to different disease situations. It provides a transparent, reproducible process and is a policy-ready tool. It has the potential to win support from Ministers and key stakeholders, whose commitment is crucial for financing and implementation of effective surveillance and control measures.

HandiResponse has been designed for studying emerging zoonotic infectious diseases, although it has the potential to be broadened to cover other diseases. The rationale for focusing on emerging zoonotic diseases and development of such a disease management program comprises:

- the importance of emerging zoonotic diseases and the potential socio-economic impacts - in particular, the effects on human populations caused by these diseases such as HPAI, Ebola Hemorrhagic Fever; and
- the challenges in predicting and managing these diseases. For instance, in the early phase of an emerging zoonotic disease outbreak, perhaps little is known about the epidemiology of the disease. Besides, many of the countries that are most at risk from such diseases do not have the skilled people to provide a sound assessment on what action is needed, how best to assess the potential impact of the disease to the country and what immediate response is appropriate (Agyepong, 2014). A software tool, such as HandiResponse, could be used to facilitate decision-making on how to respond to such risk situations and support resource-poor countries in mounting an appropriate response.

A spatial and temporal model representing the particular disease under study is critical for the development of technically and economically efficient surveillance portfolios and other appropriate disease mitigation measures. This is achieved by HandiSpread, a modified version of InterSpread Plus. HandiSpread is able to model a disease affecting animal populations, with spillover into human populations, but no significant human-to-human spread (such as avian influenza H5N1). It could be linked to models which deal specifically with transmission within human populations, for diseases which spillover into human populations and may then cause continuing human-to-human spread (such as Ebola haemorrhagic fever).

The program operates within HandiView, a platform for data management in health events which involve people and/or animals. Different computer programming languages have been used for the development of HandiResponse to maximize its functionality and user friendliness: C# was used for the webserver design, Java Script for the development of web interface, R for the development of HandiResponse report template and Visual Basic for OptiSurv module.

3.5.2. Risk mapping

3.5.2.1. Definition, typology and development

Risk mapping has a number of definitions and connotations. For disaster management, risk mapping is a process of identifying high-risk areas, done by relating the nature of a hazard, such as an earthquake, to the terrain and to the probability that an event will occur at a particular location. In public health, disease mapping refers to the geographical distribution of a disease within a population (Andrew, 2001; Lawson, Williams, & Williams, 2001). It is considered as an exploratory method of analysis used to get an impression of the spatial distribution of a disease and/or the corresponding risks. Risk maps can also be outcomes of models of disease transmission based on spatial and temporal data. These models incorporate, to varying degrees, epidemiological, entomological, climatic and environmental information (Kitron, 2000). Kitron elaborated that as a major method of spatial epidemiology, disease risk mapping could help test a hypothesis, identify gaps in our knowledge, provide a direction for surveillance and control efforts, or evaluate the actual or potential effectiveness of an intervention. There are varied examples of using risk maps for identifying which geographic locations are most suitable for disease events to occur or sustain, based upon assessment of relevant biotic and abiotic risk factors, or else of information derived directly from past disease incidents (Clements, Pfeiffer et al., 2007; Fichet-Calvet & Rogers, 2009; Glass, Cheek et al., 1995; Pigott, Bhatt et al., 2014; Snow, Craig et al., 1999).

Risk maps can be categorized in varied ways, which are summarized in Table 3-3 (Kitron, 2000; Myers, Rogers et al., 2000; Ostfeld, Glass, & Keesing, 2005).

Table 3-3. Typology of risk map

Criteria	Typology
Nature of map	•Empirical •Model based: (statistical or biological process-based) •Knowledge based
What is mapped	•Disease or pathogen •Risk factors or their covariates: abiotic and biotic factors such as environmental factors, vectors, reservoirs
When is mapped for	•Retrospective, historical pattern; •Prospective: abiotic variables showing the greatest degree of spatial concordance with vector distribution are assumed to be causative and are used to predict current distributions of vectors in unstudied areas or future distributions •Real-time: risk maps being updated, as frequently as new occurrence data are assimilated
How is mapped	•Static: snapshot of distributions of arthropod vectors, vertebrate reservoirs, or actual cases of disease in the host •Dynamic: analyzing what factors govern the spatial pattern and rate of spread of disease temporally or spatially; analyzing travelling 'waves' of epidemics
How is presented	•Dot map for point (or case-event) data •Raster map, choropleth maps for regional data •Isopleth maps for geostatistical data
How is used	•Testing hypothesis •Designing surveillance and interventions •Evaluation of impacts of disease and impacts of control programs

Although infectious diseases, in particular, zoonotic diseases, have their favourable habitats (ecological niches) (Hogerwerf, Wallace et al., 2010; Kurtenbach, Hanincová et al., 2006; Vandermeer, 1972), to define or predict their precise spatial and temporal distribution is challenging. The ideal risk map is actually the outcome of modelling disease transmission and accurately representing the ecological, biological and behavioral processes influenced by host, pathogen and surrounding environmental and climatic factors (Slingenbergh, 2004; Karesh, Dobson et al., 2012; Wilcox & Gubler, 2005). Such a process may misrepresent the dynamic process for any given disease if (i) the input data are out-of-date, or simply wrong; (ii) the explanatory variables used are inappropriate; (iii) the model itself is wrong (Hirzel, Hausser et al., 2002; D. J. Rogers & Randolph, 2003; Woolhouse, 2011). In addition, an organism or disease will not occupy all suitable habitats. Hence even by combining geographical information systems, remote sensing, computer technology and statistical tools, risk maps developed by using disease risk factors and their surrogates, individually or collectively can only be the analyst's best estimate of the likelihood of disease spread at a geographic location, and in some cases may only be relevant to a particular time period or season.

Risk maps therefore need reinforcement and validation by undertaking appropriately targeted surveillance to support or challenge the map.

3.5.2.2. Data sources for risk mapping

Data and information used for risk map development are usually generated in three ways: empirical investigation, experts' estimation and remotely sensed data (Stevens & Pfeiffer, 2011). Risk mapping can be data intensive and sometimes having adequate data from national sources for risk mapping turns out to be beyond the capacity of resource-poor countries, especially when a novel infectious disease is implicated. Utilization of global or regional remotely sensed data as the proxies or surrogates for risk factors presents the most promising and feasible option under these circumstances. Nowadays, an ever-increasing range of remotely sensed environmental data sets relevant to diseases can be accessed in the public domain or at minimal cost, and with steadily improving spatial and data scale resolution. There are some major advantages for using remotely sensed data/information: (i) free or low cost; (ii) standardized and of good quality so that system errors can be possibly minimized[3]; (iii) less chance for being modified or manipulated as a result of data protection, confidentiality legislation, politics or security reasons; (iv) the data can be used for any user-defined geographic areas without being limited by national borders; (v) the range of measured and derived variables now available goes far beyond the simple variables initially reported from the early satellites, and the range is growing rapidly. Remote-sensed data is particularly useful when predicting and analyzing disease hotspots and spread across national boundaries. However, using proxy abiotic and biotic environment factors to represent or predict disease risk or incidence is not guaranteed to be free from problems. This is especially likely when no attempt is made to validate these derived layers through field observations (ground truthing) or the verification process has been only focused on presence data of a disease and its associated risk factors (Rogers & Randolph, 2003).

To facilitate risk mapping, HandiResponse groups emerging zoonoses into different "epitypes" according to the mode of transmission (Table 3-4), following WHO's recommendations (WHO & FAO, 1967). The rationale behind this is that diseases classified under the same epitype are likely to have a similar set of risk factors and risk proxies, because they have similar transmission mechanisms.

---

[3] http://www.fao.org/geonetwork/srv/en/main.home

71

Table 3-4. Epitypes of Zoonoses

Epitype and Definition	Example
Direct Zoonoses**:** diseases are transmitted from an infected vertebrate host to another host by direct contact, fomite or mechanical vector. The pathogen does not undergo developmental change or propagation during the transmission. This group of zoonoses can be further categorized as: • Limited environmental influence on direct transmission host to host • Strong environmental influence on direct transmission host to host • Transmission heavily reliant on fomites as intermediates in transmission • Food borne • Transmission by mechanical vector • Transmission by biological vector, but the vector stage does not involve a different replication process for the agent (facultative vector)	 • Rabies • Avian Influenza • Brucellosis • BSE, E.coli (O157), Hepatitis A, E • Salmonella spp. by Musca domestica • Crimean Congo Haemorrhagic Fever, Lyme disease
Cyclozoonoses: are zoonotic diseases which require more than one vertebrate host but no invertebrate host is needed.	• Echinococcosis
Metazoonoses: are zoonotic diseases which require an invertebrate host in which the pathogen must go through a different stage before it can infect a vertebrate host (obligate vector).	• Leishmaniasis
Saprozoonoses: are diseases transferred through a non-animal reservoir, such as a plant, or through the abiotic environment, such as through water or soil	• Anthrax

It is important to abide by the rule of parsimony or non-redundancy for the selection of risk factors and their proxies for a given disease when mapping a disease. This means a minimum set of relevant risk factors and proxies that are either un-correlated or have a low level of correlation would be selected for a given disease risk mapping. Further criteria for the selection of risk factors and their proxies include (i) epidemiological relevance: the selected factors need to be epidemiological relevant, either based on evidence or informed by experts' opinion; (ii) completeness: all the relevant risk components, factors/proxies are included; (iii) measurability: a numeric value can be assigned to a risk factor/proxy selected (Malczewski, 2002). A summary of possible risk layers is illustrated in Table 3-5 (Molesworth, Thomson et al., 2002).

Table 3-5. Categorization of risk factors related to emerging infectious diseases

Mechanism	Component	Example Factor/Proxy
Susceptibility	Human	Density Sub group: age; percent of rural population, or poor population, etc. Co-location and mixing of different susceptible species Existence of preventive measures such as vaccination
	Animal(s)	Density Sub-group: age, sex
	Vector	Density
Exposure to pathogen(s)	Movement	Movement of hosts and contaminated products Distance to road, waterway, or road density, etc.
Pathogen viability	Environment	Temperature Altitude Other environment factors such as % of cropland, forest/shrub land/grassland, rice paddy
	Climate	Rainfall

3.5.2.3. Risk factor scoring, standardization and weight elicitation

HandiMap employs weighted summation (WSum) approach for integrating data on all selected risk factor layers for risk map development. WSum is a Multicriteria Decision Analysis (MCDA) method that has been widely used. It can be implemented within a GIS environment and is fairly unchallenging for decision makers to understand. One typical example of such an approach is the Weighted Linear Combination (WLC) (Jiang and Eastman 2000; Ayalew, Yamagishi et al. 2004; Malczewski, 2006; Malczewski, 2011; Shahabi, Keihanfard et al. 2014). WSum is very suitable for participatory and transparent decision process. Some GIS systems such as IDRISI have built-in routines for WLC. Two of the most frequently cited weaknesses of the method include (i) trade-off or substitutability which means a low score on one risk factor can be compensated by a high score on another and (ii) requirement of independence between selected risk factors which may be difficult to be fulfilled (Drobne & Lisec, 2009).

3.5.2.4. Standardization

When using environmental, climatological or other data to represent or estimate disease risk, the raw metrics of each risk layer need to be converted to risk scores representing the probability of a disease occurrence. Besides, the measured values on all risk layers need to be transformed into a mutually compatible scale through standardization. HandiResponse accommodates flexible approaches for standardization. Users can choose linear scale transformation methods such as maximum standardization, interval standardization and goal standardization, and can also use a non−linear value function approach. An underlying relationship between particular variables and disease risk needs to be assumed. Selection of an appropriate function that relates the variable to disease risk could be

informed by literature review, empirical studies or expert's opinion. A sample of selected risk factors and their proposed relationship with HPAI is provided in Table 3-6 (Gilbert & Pfeiffer, 2012). This information will be used in Chapter 5. It is worth noting that the same risk factor might be proposed by different authors as positively or negatively associated with the disease under study. Besides, caution needs to be practiced when extrapolating findings from one geographical/ecological setting to another.

Table 3-6. Selected key risk factors, environmental predictors and their proposed relationship with HPAI H5N1 outbreaks in South East Asia

Risk Factors	Relationship	References
Density of ducks	positive, particularly the density of free grazing ducks	(Sturm-Ramirez, Hulse-Post et al. 2005; Hulse-Post, Sturm-Ramirez et al. 2005; Gilbert, Chaitaweesup et al. 2006; Pfeiffer Minh et al. 2007; Tiensin, Ahmed et al. 2009; Hogerwerf, Wallace et al. 2010; Gilbert, Newman et al. 2010)
Proportion of land used for rice/rice cropping density	Positive association with duck density; Negative association with chicken density	(Tiensin, Chaitaweesup et al. 2005; Gilbert, Chaitaweesup et al. 2006; Edan & Bourgeois, 2006; Pfeiffer, Minh et al. 2007; Gilbert, Xiao et al. 2008; Paul, Tavornpanich et al. 2010; Hogerwerf, Wallace et al. 2010; Loth, Gilbert et al. 2011; Gilbert, Xiao et al. 2007)
Distance to highway	positive association with poultry farming/trading; elevated risk is associated the distance within 10 km to any highways	(Gautheir-Clerc, Lebarbenchon et al. 2007; Vannier, 2007; Ward, Maftei et al. 2008; Paul, Tavornpanich et al. 2010; Cao, Xu et al. 2010; Ge, Haining et al. 2012)
Elevation	positive or negative association with rice cropping, duck density; low elevation is positively associated with occurrence of HPAI in wild birds	(Gilbert, Chaitaweesup et al. 2006; Pfeiffer, Minh et al. 2007; Gilbert, Xiao et al. 2008; Tiensin, Ahmed et al. 2009; Paul, Tavornpanich et al. 2011; Si, Wang et al. 2010; Takekawa, Newman et al. 2010; Loth, Gilbert et al. 2011)
Density of human population	positive (rural pop) or negative (city); medium density is positively associated with elevated HPAI risk	(Gilbert, Chaitaweesup et al. 2006, 2008; Tiensin, Ahmed et al. 2009; Loth, Gilbert et al. 2011, Yuniana, de Vlas et al. 2010; Paul, Tavornpanich et al. 2010; Hogerwerf, Wallace et al. 2010; Martin, Pfeiffer et al. 2011)

3.5.2.5. Weight elicitation

The effect of selected risk factors and their proxies on disease occurrence varies. Some risk factors are epidemiologically more important than others, a small change in value of them (steep slope) could cause a substantial decrease or increase in disease occurrence. This means such factors carry more weight than others.

A variety of subjective, objective and integrated weight elicitation methods exist (Drobne & Lisec, 2009; Ma, Fan, & Huang, 1999; Roszkowska, 2013). There has been no "best method" but only the "most suitable" one for weight elicitation. It is critical to test the reliability and validity of the chosen method(s) by conducting sensitivity analysis (Bertsch & Geldermann, 2008). A number of commonly used ones include (i) ranking methods such as rank order centroid (ROC) (Stillwell, Seaver, & Edwards, 1981); (ii) rating methods such as direct rating (DR) and point allocation (PA) (Bottomley, Doyle, & Green, 2000; Doyle, Green, & Bottomley, 1997) ; (iii) pairwise comparison methods such as analytical hierarchy process (AHP) (Buede, 1992; Saaty, 1986, 1987, 1990, 2013); (iv) trade-off analysis methods such as swing weights technique (Edwards & Barron, 1994) and (v) objective methods such as entropy and standardized deviation method (Diakoulaki, Mavrotas, & Papayannakis, 1995; Yoon & Hwang, 1995). AHP will be used in Chapter 4 for facilitating weight elicitation in studying HPAI (H5N1) in Vietnam and in Chapter 5 for studying Crimean Congo Haemorrhagic Fever (CCHF). The advantages of using such an approach include (i) it allows a transparent and participatory process; (ii) the method incorporates consistency check and (iii) it is an appropriate method to be used in conjunction with GIS. Besides, a number of free or commercial software products to apply AHP are available to facilitate implementation of such an approach (Ossadnik & Lange, 1999). The AHPVEC function in SAS/IML® can also perform AHP (Alexander, 2012). In the study on HPAIV H5N1 in southern Vietnam, we attempted different methods and the results are presented below (Figure 3-3).
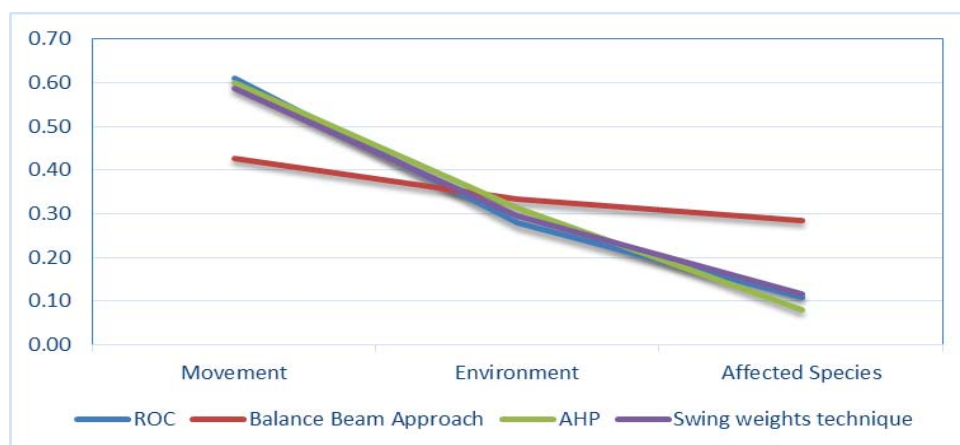
Figure 3-3. Results of weight elicitation by different methods
Note: Y-axis is weight value with a range from 0 to 1. The higher the weight, the larger the value; ROC stands for Rank Order Centroid; AHP stands for Analytical Hierarchy Process.

In HandiResponse, weight elicitation is implemented in two steps: (i) weight is assigned to risk components/groups; (ii) weight is assigned to risk factors. Such a hierarchical arrangement has been used because it is particularly useful to handle complexity and when dealing with a large number of risk factors (Figure 3-4).
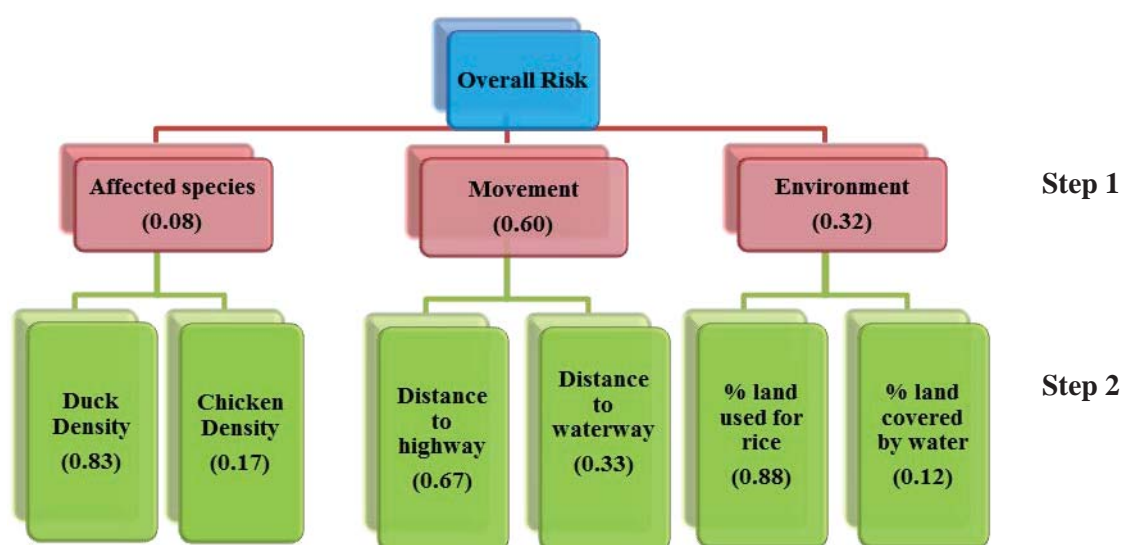
Figure 3-4. Example: AHP result of weight elicitation by risk component and layer for HPAI in southern Vietnam

3.5.3.Models for disease introduction and spread

Developing a best fit spatial and temporal model is essential for studying disease behaviour hence identifying and estimating surveillance and other disease control strategies. Models have been long used for representing disease behaviours predictively, retrospectively and in real time (Garner Lebarbenchon, & Thomas 2007; Taylor, 2003). The models for representing the relationship between risk (or risk surrogate) and disease generally fall into two categories. The first is the statistical approach (which is based on an assumption that a statistical relationship exists between past case counts and risk factors, plus environmental predictors) (Martin, Pfeiffer, et al., 2011; Randolph, 2001). This approach is typically applied by deterministic mathematical modelling. On the other hand, the biological process-based approach attempts to capture the biology of transmission processes (Focks, Daniels et al., 1995), and typically uses stochastic simulation. Using the statistical approach for prediction of future disease occurrence or extrapolation to unstudied areas makes the assumption that the disease in the future or in the unstudied areas behaves the same as in the previously investigated situations. Whereas using the biological approach requires details on all the important parameters, and on the relationships between them and disease occurrence.

Which model to choose is a technical as well as a practical question. In the absence of full knowledge of transmission pathways for a disease, the statistical approach can be used as a temporary substitute for the biological process-based approach (Myers, Rogers et al., 2000).

For modelling the HPAI H5N1 outbreak in South Vietnam, we have used HandiSpread, a stochastic disease modelling program which uses the biological process-based approach.

3.5.4. Optimization of surveillance program

Surveillance is an important first step for disease control. Achieving higher level of efficiency and accountability is the ultimate goal for all disease surveillance and control programs. Presenting the economic case for a surveillance program could be a way to break the vicious cycle such as chronic under-funding and suboptimal performance of the system. For governments and increasingly for private sector organizations which commission and allocate resources for surveillance activities, one of the key objectives is to achieve best performance within a given allocation, or the lowest investment required to achieve predefined objectives.

Several useful studies have attempted to develop methods for optimizing technical performance of surveillance systems. For instance, by using a stochastic scenario tree model of the vector-borne disease bluetongue, Hadorn *et al.* simulated the sensitivity (median and 95% confidence interval) of each individual surveillance system component, and combined surveillance strategies. The comparative cost effectiveness study revealed that the surveillance approach combining passive clinical surveillance in cattle and sheep and targeted bulk milk testing in cattle herds located in high-risk areas could achieve the highest sensitivity, though with a slightly increased cost (Hadorn, Racloz et al., 2009). The same approach was also used for opimizing the HPAI H5N1 surveillance program in Thailand (Goutard, Tavornpanich et al., 2012). However, these studies failed to factor in the economic impacts of the disease due to productivity loss and possible animal trade bans, etc., in deciding on optimal surveillance options. Prattley *et al.* explored a novel method for optimizing resource allocation for risk-based surveillance program in New Zealand. By using portfolio theory, they demonstrated how to allocate available resources among the surveillance programs for different exotic diseases, and allocate (surveillance) resource by region and month according perceived relative risk levels. Their argument was that if a particular disease spread unevenly by space, population and time, higher risk strata would deserve more resource allocation because of "high return on investment". However, higher standard deviation of return was associated with high risk strata. Two variants of portfolio theory, risk aversion (safety first model) or risk taking (single index model to achieve efficiency frontier) were explored. The study revealed that both risk-based approaches, in particular, the safety first approach, would perform better than the conventional proportional sampling approach (Prattley, Morris et al., 2007). Such a method is useful when only a single surveillance component is to be implemented. However in the real world, a disease surveillance program seldom comprises only one component. Other useful surveillance optimization measures practiced for early disease/event detection include using maximum coverage model for searching surveillance sites, or implementing different searching or analytical algorithms for aberration detection, and stratifying data and adding

new data sources (Hutwagner, Browne et al., 2005; Polgreen, Chen et al., 2009; Scarpino, Dimitrov, & Meyers, 2012; Sparks, 2013; Wang, Zeng et al., 2010; Zhang, Jamal et al., 2011).

Unfortunately, few surveillance evaluation activities have been implemented by focusing on cost effectiveness and cost/benefit even though they are highly recommended (Drewe, Hoinville et al., 2012; German, Lee et al., 2001; Hoinville, Alban et al., 2013). There has been no strategy like OptiSurv that incoporates the concept of multi-criteria decision making, not only identifying the best surveillance portfolios with the earliest detection of disease but also the most economically efficient ones, which maximize net economic benefit of the surveillance activity.

3.6. Conclusion

HandiResponse, combining risk mapping, infectious disease modelling, surveillance program planning and optimization in one package, has been designed to support disease surveillance activities, particularly for emerging or globally spreading diseases in the first instance. It can potentially facilitate effective communication between technical staff and policy makers on disease surveillance and improve technical and economic efficiency of the programs for disease surveillance, in particular, those in developing world. Subsequent chapters will demonstrate its operation, and provide examples of its use.

3.7. References

Agyepong, L. (2014). "A systems view and lessons from the ongoing Ebola Virus Disease (EVD) outbreak in West Africa." Ghana medical journal **48**(3): 168-172.

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

Alexander, M. (2012). "Decision-Making using the Analytic Hierarchy Process (AHP) and SAS/IML'." Available from (Last checked 28th July 2014).

Andrew, B., Lawson, Fiona, L.R., Williams (2001). An introductory guide to disease mapping. West Sussex, UK, John Wiley and Sons Ltd.

Bertsch, V. and J. Geldermann (2008). "Preference elicitation and sensitivity analysis in multicriteria group decision support for industrial risk and emergency management." International Journal of Emergency Management **5**(1-2): 7-24.

Bottomley, P. A., J. R. Doyle and R. H. Green (2000). "Testing the reliability of weight elicitation methods: direct rating versus point allocation." Journal of Marketing Research **37**(4): 508-513.

Buede, D. M. (1992). "Software Review-Three Packages for AHP: Criterium, Expert Choice and HIPRE 3 +." Journal of Multi-Criteria Decision Analysis **1**: 119-121.

Calvo-Artavia, F. F., L. R. Nielsen and L. Alban (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.

Cao, C., M. Xu, C. Chang, Y. Xue, S. Zhong, L. Fang, W. Cao, H. Zhang, M. Gao and Q. He (2010). "Risk analysis for the highly pathogenic avian influenza in Mainland China using meta-modeling." Chinese Science Bulletin 55(36): 4168-4178.

Carroll, L. N., A. P. Au, L. T. Detwiler, T.-c. Fu, I. S. Painter and N. F. Abernethy (2014). "Visualization and analytics tools for infectious disease epidemiology: A systematic review." Journal of biomedical informatics **51**: 287-298.

Christensen, J., F. El Allaki and A. Vallières (2014). "Adapting a scenario tree model for freedom from disease as surveillance progresses: the Canadian notifiable avian influenza model." Preventive Veterinary Medicine **114**(2): 132-144.

Clements, A. C. A., D. U. Pfeiffer, V. Martin, C. Pittiglio, N. Best and Y. Thiongane (2007). "Spatial risk assessment of Rift Valley fever in Senegal." Vector-Borne and Zoonotic Diseases **7**(2): 203-216.

Diakoulaki, D., G. Mavrotas and L. Papayannakis (1995). "Determining objective weights in multiple criteria problems: the CRITIC method." Computers & Operations Research **22**(7): 763-770.

Doyle, J. R., R. H. Green and P. A. Bottomley (1997). "Judging relative importance: direct rating and point allocation are not equivalent." Organizational behavior and human decision processes **70**(1): 65-72.

Drewe, J., L. Hoinville, A. Cook, T. Floyd and K. Stärk (2012). "Evaluation of animal and public health surveillance systems: a systematic review." Epidemiology and infection **140**(04): 575-590.

Drobne, S. and A. Lisec (2009). "Multi-attribute decision analysis in GIS: weighted linear combination and ordered weighted averaging." Informatica **33**(4).

Edan, M. and N. Bourgeois (2006). "Review of free-range duck farming systems in Northern Vietnam and assessment of their implication in the spreading of the highly pathogenic (H5N1) strain of avian influenza (HPAI)." Lyon, France: Agronomes et Veterinaires sans Frontieres.

Edwards, W. and F. H. Barron (1994). "SMARTS and SMARTER: Improved simple methods for multiattribute utility measurement." Organizational behavior and human decision processes **60**(3): 306-325.

Fichet-Calvet, E. and D. J. Rogers (2009). "Risk Maps of Lassa Fever in West Africa." Plos Neglected Tropical Diseases **3**(3).

Focks, D. A., E. Daniels, D. G. Haile and J. E. Keesling (1995). "A simulation model of the epidemiology of urban dengue fever: literature analysis, model development, preliminary validation, and samples of simulation results." American Journal of Tropical Medicine and Hygiene **53**(5): 489-506.

Garner, M. G., C. Dubé, M. A. Stevenson, R. L. Sanson, C. Estrada and J. Griffin (2007). "Evaluating alternative approaches to managing animal disease outbreaks–the role of modelling in policy formulation." Vet. ital **43**(2): 285-298.

Gauthier-Clerc, M., C. Lebarbenchon and F. Thomas (2007). "Recent expansion of highly pathogenic avian influenza H5N1: a critical review." Ibis 149(2): 202-214.

Ge, E., R. Haining, C. P. Li, Z. Yu, M. Y. Waye, K. H. Chu and Y. Leung (2012). "Using knowledge fusion to analyze avian influenza H5N1 in East and Southeast Asia."

Geanuracos, C. G., S. D. Cunningham, G. Weiss, D. Forte, L. M. H. Reid and J. M. Ellen (2007). "Use of geographic information systems for planning HIV prevention interventions for high-risk youths." American Journal of Public Health **97**(11): 1974-1981.

German, R. R., L. Lee, J. Horan, R. Milstein, C. Pertowski and M. Waller (2001). "Updated guidelines for evaluating public health surveillance systems." MMWR Recomm Rep **50**(1-35).

Gilbert, M., P. Chaitaweesup, T. Parakamawongsa, S. Premashthira, T. Tiensin, W. Kalpravidh, H. Wagner and J. Slingenbergh (2006). "Free-grazing ducks and highly pathogenic avian influenza, Thailand." Emerging infectious diseases 12(2): 227-234.

Gilbert, M., S. H. Newman, J. Y. Takekawa, L. Loth, C. Biradar, D. J. Prosser, S. Balachandran, M. V. S. Rao, T. Mundkur and B. Yan (2010). "Flying over an infected landscape: distribution of highly pathogenic avian influenza H5N1 risk in South Asia and satellite tracking of wild waterfowl." EcoHealth 7(4): 448-458.

Gilbert, M. and D. U. Pfeiffer (2012). "Risk factor modelling of the spatio-temporal patterns of highly pathogenic avian influenza (HPAIV) H5N1: a review." Spatial and spatio-temporal epidemiology **3**(3): 173-183.

Gilbert, M., X. Xiao, P. Chaitaweesub, W. Kalpravidh, S. Premashthira, S. Boles and J. Slingenbergh (2007). "Avian influenza, domestic ducks and rice agriculture in Thailand." Agriculture, ecosystems & environment 119(3): 409-415.

Gilbert, M., X. Xiao, D. U. Pfeiffer, M. Epprecht, S. Boles, C. Czarnecki, P. Chaitaweesub, W. Kalpravidh, P. Q. Minh and M. J. Otte (2008). "Mapping H5N1 highly pathogenic avian influenza risk in Southeast Asia." Proceedings of the National Academy of Sciences 105(12): 4769-4774.

Glass, G. E., B. S. Schwartz, J. M. Morgan III, D. T. Johnson, P. M. Noy and E. Israel (1995). "Environmental risk factors for Lyme disease identified with geographic information systems." American Journal of Public Health **85**(7): 944-948.

Goutard, F. L., M. Paul, S. Tavornpanich, I. Houisse, K. Chanachai, W. Thanapongtharm, A. Cameron, K. D. Stärk and F. Roger (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.

Hadorn, D. C., V. Racloz, H. Schwermer and K. D. Stärk (2009). "Establishing a cost-effective national surveillance system for Bluetongue using scenario tree modelling." Veterinary research **40**(6): 1-14.

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

Hirzel, A. H., J. Hausser, D. Chessel and N. Perrin (2002). "Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data?" Ecology **83**(7): 2027-2036.

Hogerwerf, L., R. G. Wallace, D. Ottaviani, J. Slingenbergh, D. Prosser, L. Bergmann and M. Gilbert (2010). "Persistence of highly pathogenic avian influenza H5N1 virus defined by agro-ecological niche." EcoHealth **7**(2): 213-225.

Hoinville, L., L. Alban, J. Drewe, J. Gibbens, L. Gustafson, B. Häsler, C. Saegerman, M. Salman and K. Stärk (2013). "Proposed terms and concepts for describing and evaluating animal-health surveillance systems." Preventive veterinary medicine **112**(1): 1-12.

Hulse-Post, D., K. Sturm-Ramirez, J. Humberd, P. Seiler, E. Govorkova, S. Krauss, C. Scholtissek, P. Puthavathana, C. Buranathai and T. Nguyen (2005). "Role of domestic ducks in the propagation and biological evolution of highly pathogenic H5N1 influenza viruses in Asia." Proceedings of the National Academy of Sciences of the United States of America 102(30): 10682-10687.

Hutwagner, L., T. Browne, G. M. Seeman and A. T. Fleischauer (2005). "Comparing aberration detection methods with simulated data." Emerg Infect Dis **11**(2): 314-316.

Slingenbergh, M. G., K. de Balogh and W. Wint (2004). "Ecological sources of zoonotic diseases." Rev. sci. tech. Off. int. Epiz., **23**(2): 467-484.

Kahn, L. H. (2006). "Confronting zoonoses, linking human and veterinary medicine." Emerging Infectious Diseases **12**(4): 556.

Kalluri, S., P. Gilruth, D. Rogers and M. Szczur (2007). "Surveillance of arthropod vector-borne infectious diseases using remote sensing techniques: a review." PLoS Pathog **3**(10): 1361-1371.

Karesh, W. B., A. Dobson, J. O. Lloyd-Smith, J. Lubroth, M. A. Dixon, M. Bennett, S. Aldrich, T. Harrington, P. Formenty and E. H. Loh (2012). "Ecology of zoonoses: natural and unnatural histories." The Lancet **380**(9857): 1936-1945.

Kitron, U. (2000). "Risk maps: transmission and burden of vector-borne diseases." Parasitol Today **16**(8): 324-325.

Kuiken, T., F. Leighton, R. Fouchier and J. LeDuc (2005). "Pathogen surveillance in animals." Science **309**(5741): 1680.

Kulasekera, V. L., L. Kramer, R. S. Nasci, F. Mostashari, B. Cherry, S. C. Trock, C. Glaser and J. R. Miller (2001). "West Nile virus infection in mosquitoes, birds, horses, and humans, Staten Island, New York, 2000." Emerging infectious diseases **7**(4): 722.

Kurtenbach, K., K. Hanincová, J. I. Tsao, G. Margos, D. Fish and N. H. Ogden (2006). "Fundamental processes in the evolutionary ecology of Lyme borreliosis." Nature Reviews Microbiology **4**(9): 660-669.

Lawson, A. B., F. L. Williams and F. Williams (2001). An introductory guide to disease mapping, John Wiley.

Loth, L., M. Gilbert, J. Wu, C. Czarnecki, M. Hidayat and X. Xiao (2011). "Identifying risk factors of highly pathogenic avian influenza (H5N1 subtype) in Indonesia." Preventive veterinary medicine 102(1): 50-58.

Ma, J., Z.-P. Fan and L.-H. Huang (1999). "A subjective and objective integrated approach to determine attribute weights." European journal of operational research **112**(2): 397-404.

Malczewski, J. (2002). "On the use of weighted linear combination method in GIS: common and best practice approaches." Transactions in GIS **4**(1): 5-22.

Martin, P., A. Cameron, K. Barfod, E. Sergeant and M. Greiner (2007). "Demonstrating freedom from disease using multiple complex data sources: 2: Case study—Classical swine fever in Denmark." Preventive Veterinary Medicine **79**(2): 98-115.

Martin, V., D. U. Pfeiffer, X. Zhou, X. Xiao, D. J. Prosser, F. Guo and M. Gilbert (2011). "Spatial distribution and risk factors of highly pathogenic avian influenza (HPAI) H5N1 in China." PLoS Pathog **7**(3): e1001308.

Molesworth, A. M., M. C. Thomson, S. J. Connor, M. P. Cresswell, A. P. Morse, P. Shears, C. A. Hart and L. E. Cuevas (2002). "Where is the meningitis belt? Defining an area at risk of epidemic meningitis in Africa." Trans R Soc Trop Med Hyg **96**(12174770): 242-249.

Myers, M. F., D. Rogers, J. Cox, A. Flahault and S. Hay (2000). "Forecasting disease risk for increased epidemic preparedness in public health." Advances in parasitology **47**: 309-330.

Ortiz-Pelaez, A., D. U. Pfeiffer, R. J. Soares-Magalhaes and F. J. Guitian (2006). "Use of social network analysis to characterize the pattern of animal movements in the initial phases of the 2001 foot and mouth disease (FMD) epidemic in the UK." Preventive Veterinary Medicine **76**(1-2): 40-55.

Ossadnik, W. and O. Lange (1999). "AHP-based evaluation of AHP-Software." European journal of operational research **118**(3): 578-588.

Ostfeld, R. S., G. E. Glass and F. Keesing (2005). "Spatial epidemiology: an emerging (or re-emerging) discipline." Trends in Ecology & Evolution **20**(6): 328-336.

Paul, M., S. Tavornpanich, D. Abrial, P. Gasqui, M. Charras-Garrido, W. Thanapongtharm, X. Xiao, M. Gilbert, F. Roger and C. Ducrot (2010). "Anthropogenic factors and the risk of highly pathogenic avian influenza H5N1: prospects from a spatial-based model." Veterinary research 41(3): 28.

Pfeiffer, D. U., P. Q. Minh, V. Martin, M. Epprecht and M. J. Otte (2007). "An analysis of the spatial and temporal patterns of highly pathogenic avian influenza occurrence in Vietnam using national surveillance data." The Veterinary Journal 174(2): 302-309.

Pigott, D. M., N. Golding, A. Mylne, Z. Huang, A. J. Henry, D. J. Weiss, O. J. Brady, M. U. Kraemer, D. L. Smith and C. L. Moyes (2014). "Mapping the zoonotic niche of Ebola virus disease in Africa." Elife **3**: e04395.

Polgreen, P. M., Z. Chen, A. M. Segre, M. L. Harris, M. A. Pentella and G. Rushton (2009). "Optimizing influenza sentinel surveillance at the state level." American journal of epidemiology: kwp270.

Prattley, D., R. Morris, M. Stevenson and R. Thornton (2007). "Application of portfolio theory to risk-based allocation of surveillance resources in animal populations." Preventive veterinary medicine **81**(1): 56-69.

Presi, P., K. D. C. Staerk, L. Knopf, E. Breidenbach, M. Sanaa, J. Frey and G. Regula (2008). "Efficiency of risk-based vs. random sampling for the monitoring of tetracycline residues in slaughtered calves in Switzerland." Food Additives and Contaminants **25**(5): 566-573.

Qing, T., M. Saijo, H. Lei, M. Niikura, A. Maeda, T. Ikegami, W. Xinjung, I. Kurane and S. Morikawa (2003). "Detection of immunoglobulin G to Crimean-Congo hemorrhagic fever virus in sheep sera by recombinant nucleoprotein-based enzyme-linked immunosorbent and immunofluorescence assays." Journal of virological methods **108**(1): 111-116.

Racloz, V., G. Venter, C. Griot and K. D. C. Stark (2008). "Estimating the temporal and spatial risk of bluetongue related to the incursion of infected vectors into Switzerland." BMC Veterinary Research **4**(42): (15 October 2008)-(2015 October 2008).

Randolph, S. E. (2001). "The shifting landscape of tick-borne zoonoses: tick-borne encephalitis and Lyme borreliosis in Europe." Philosophical Transactions of the Royal Society B: Biological Sciences **356**(1411): 1045-1056.

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

Rogers, D. J. and S. E. Randolph (2003). "Studying the global distribution of infectious diseases using GIS and RS." Nature Reviews Microbiology **1**(3): 231-237.

Roszkowska, E. (2013). "Rank Ordering Criteria Weighting Methods–a Comparative Overview."

Saaty, R. W. (1987). "THE ANALYTIC HIERARCHY PROCESS - WHAT IT IS AND HOW IT IS USED." Mathematical Modelling **9**(3-5): 161-176.

Saaty, T. L. (1986). "A NOTE ON THE AHP AND EXPECTED VALUE THEORY." Socio-Economic Planning Sciences **20**(6): 397-398.

Saaty, T. L. (1990). "HOW TO MAKE A DECISION - THE ANALYTIC HIERARCHY PROCESS." European Journal of Operational Research **48**(1): 9-26.

Saaty, T. L. (2013). "The Modern Science of Multicriteria Decision Making and Its Practical Applications: The AHP/ANP Approach." Operations Research **61**(5): 1101-1118.

Scarpino, S. V., N. B. Dimitrov and L. A. Meyers (2012). "Optimizing provider recruitment for influenza surveillance networks." PLoS Comput Biol **8**(4): e1002472.

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

Si, Y., T. Wang, A. K. Skidmore, W. F. de Boer, L. Li and H. H. Prins (2010). "Environmental factors influencing the spread of the highly pathogenic avian influenza H5N1 virus in wild birds in Europe." Ecol. Soc 15: 26.

Smieszek, T. and M. Salathé (2013). "A low-cost method to assess the epidemiological importance of individuals in controlling infectious disease outbreaks." BMC medicine **11**(1): 35.

Snow, R., M. Craig, U. Deichmann and D. Le Sueur (1999). "A preliminary continental risk map for malaria mortality among African children." Parasitology today **15**(3): 99-104.

Sparks, R. (2013). "Challenges in designing a disease surveillance plan: What we have and what we need?" IIE Transactions on Healthcare Systems Engineering **3**(3): 181-192.

Stark, K. D. C., G. Regula, J. Hernandez, L. Knopf, K. Fuchs, R. S. Morris and P. Davies (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): (28 February 2006)-(2028 February 2006).

Stevens, K. B. and D. U. Pfeiffer (2011). "Spatial modelling of disease using data-and knowledge-driven approaches." Spatial and spatio-temporal epidemiology **2**(3): 125-133.

Stillwell, W. G., D. A. Seaver and W. Edwards (1981). "A comparison of weight approximation techniques in multiattribute utility decision making." Organizational behavior and human performance **28**(1): 62-77.

Sturm-Ramirez, K., D. Hulse-Post, E. Govorkova, J. Humberd, P. Seiler, P. Puthavathana, C. Buranathai, T. Nguyen, A. Chaisingh and H. Long (2005). "Are ducks contributing to the endemicity of highly pathogenic H5N1 influenza virus in Asia?" Journal of virology 79(17): 11269-11279.

Takekawa, J. Y., S. H. Newman, X. Xiao, D. J. Prosser, K. A. Spragens, E. C. Palm, B. Yan, T. Li, F. Lei and D. Zhao (2010). "Migration of waterfowl in the East Asian flyway and spatial relationship to HPAI H5N1 outbreaks." Avian diseases 54(s1): 466-476.

Tavornpanich, S., I. A. Gardner, T. E. Carpenter, W. O. Johnson and R. J. Anderson (2006). "Evaluation of cost-effectiveness of targeted sampling methods for detection of Mycobacterium aviumsubsp paratuberculosis infection in dairy herds." American journal of veterinary research **67**(5): 821-828.

Taylor, N. (2003). "Review of the use of models in informing disease control policy development and adjustment." A report for Defra. Defra, London **94**.

Tiensin, T., S. S. U. Ahmed, S. Rojanasthien, T. Songserm, P. Ratanakorn, K. Chaichoun, W. Kalpravidh, S. Wongkasemjit, T. Patchimasiri and K. Chanachai (2009). "Ecologic risk factor investigation of clusters of avian influenza A (H5N1) virus infection in Thailand." Journal of Infectious Diseases 199(12): 1735-1743.

Tiensin, T., P. Chaitaweesup, T. Songserm, A. Chaising, W. Hoonsuwan, C. Buranathai, T. Parakamawongsa, S. Premashthira, A. Amonsin and M. Gilbert (2005). "Highly pathogenic avian influenza H5N1." Emerging infectious diseases 11: 1664-1672.

Tsai, P., K. A. Scott, M. Pappaioanou, M. C. Gonzalez and G. T. Keusch (2009). Sustaining global surveillance and response to emerging zoonotic diseases, National Academies Press.

Vandermeer, J. H. (1972). "Niche theory." Annual Review of Ecology and Systematics: 107-132.

Vannier, P. (2007). "Avian influenza; routes of transmission: lessons and thoughts drawn out of the past and present situation in the world and in the European Union." Pathologie Biologie 55(6): 273-276.

Wang, X., D. Zeng, H. Seale, S. Li, H. Cheng, R. Luan, X. He, X. Pang, X. Dou and Q. Wang (2010). "Comparing early outbreak detection algorithms based on their optimized parameter values." Journal of biomedical informatics **43**(1): 97-103.

Ward, M. P. (2007). "Geographic information system-based avian influenza surveillance systems for village poultry in Romania." Veterinaria italiana 43(3): 483-489.

Weinberg, M., S. Waterman, C. A. Lucas, V. C. Falcon, P. K. Morales, L. A. Lopez, C. Peter, A. E. Gutiérrez, E. R. Gonzalez and A. Flisser (2003). "The US-Mexico border infectious disease surveillance project: Establishing binational border surveillance." Emerging Infectious Diseases **9**(1): 97.

WHO, FAO. (1967). Joint FAO/WHO Committee on Zoonoses-Third Report, Geneva, Switzerland.

Wilcox, B. A. and D. J. Gubler (2005). "Disease ecology and the global emergence of zoonotic pathogens." Environmental Health and Preventive Medicine **10**(5): 263-272.

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

Woolhouse, M. (2011). "How to make predictions about future infectious disease risks." Philosophical Transactions of the Royal Society B: Biological Sciences **366**(1573): 2045-2054.

Yoon, K. P. and C.-L. Hwang (1995). Multiple attribute decision making: an introduction, Sage publications.

Yupiana, Y., S. J. de Vlas, N. M. Adnan and J. H. Richardus (2010). "Risk factors of poultry outbreaks and human cases of H5N1 avian influenza virus infection in West Java Province, Indonesia." International Journal of Infectious Diseases 14(9): e800-e805.

Zhang, L., I. Jamal, Y. Qu and G. Zeng (2011). "Optimization of event detection methods for disease surveillance." Épidémiologie et Santé Animale(59/60): 196-198.

4. Development of a generic system for creating a digital disease risk landscape – HandiMap

4.1. Abstract

Background*:* Disease risk mapping is a valuable tool in the development of epidemiologically appropriate surveillance and control efforts. However, most published mapping activities have been directed towards improved scientific understanding, rather than as practical decision-making tools. There is a need for a flexible method which can be used by countries to design surveillance and control policies for a range of diseases which are of national importance. To facilitate disease risk mapping, Human and Animal Disease Mapping tool (HandiMap) was developed.

Method*:* A system for using available spatial epidemiological information such as remote-sensed environmental and climatological variables to generate risk landscapes for any disease of concern in any country in the world has been developed. As a first illustration of the use of HandiMap, risk maps for Highly Pathogenic Avian Influenza (HPAI) H5N1 occurrence in southern Vietnam were developed. The mapping process was elaborated through a step-wise approach. The results were compared to the actual HPAI H5N1 epidemic in southern Vietnam between late 2004 and early 2005.

Results: Three types of HPAI H5N1 risk maps were successfully developed for southern Vietnam and were used to adjust transmission probabilities for this disease in the HandiSpread disease model according to variation in the height of the risk landscape.

Conclusion: HandiMap is a convenient tool for estimating and visualizing spatial variation in the risk of occurrence for a zoonotic disease in any country, and is applicable in resource and data sparse settings. It is suitable for use by epidemiologists to facilitate effective communication on infectious diseases, and in conjunction with the other modules of HandiResponse to support epidemiologically informed decision making for surveillance and other disease interventions.

4.2. Introduction

HandiResponse has been designed for (i) visualizing the disease risk landscape and representing spatial variation in the expected occurrence of a zoonotic disease both quantitatively and visually; (ii) evaluating economic benefit and costs of a single surveillance activity or a multi-component portfolio; (iii) identifying optimal use of resources for surveillance. The program comprises four modules: (i) risk map development – HandiMap; (ii) surveillance portfolio development – HandiSurv; (iii) economic impact assessment – HandiEcon and (iv) surveillance optimization – OptiSurv. HandiResponse operates within the Integrated Real-Time Information System (IRIS), a web-based data management platform supporting human and animal health. The particular implementation of IRIS used here is called HandiView.

In this chapter the operation of HandiMap to develop a risk landscape for design of surveillance portfolios is illustrated, using the example of HPAI H5N1 in southern Vietnam.

4.3. Method

4.3.1. Login to HandiResponse

To use HandiResponse, it is necessary to log into IRIS. IRIS can be accessed at http://demo.episoft.co.nz. With approval from a system administrator, a user can set up a user ID and a password to access the system (Figure 4-1). IRIS can operate in any language for which the translation of screen information has been completed, so that it can support any country-specific disease management programs in national languages. Currently it is available to operate in English, Chinese, Mongolian and Vietnamese. Login details determine a user's rights, such as the contents of the databases that the user can access and the functions that the user is authorised to perform on the data.
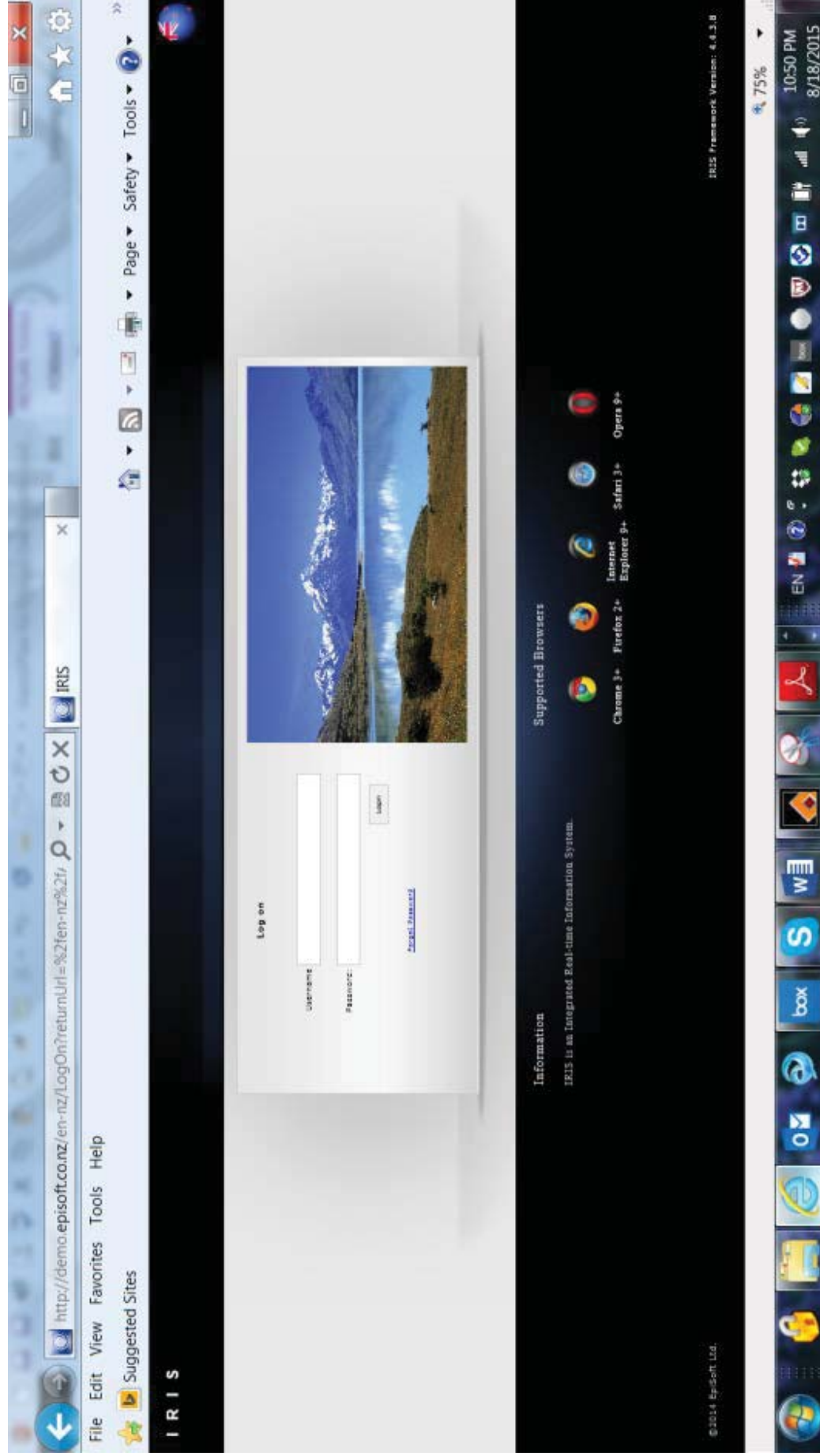
Figure 4-1. Login to IRIS

After entering IRIS, the user will see a list of specific data sets and disease management programs that can be accessed. These are called "Projects" by IRIS. Clicking on "HandiResponse" will allow the user to enter into HandiResponse (Figure 4-2. Entering HandiResponse).
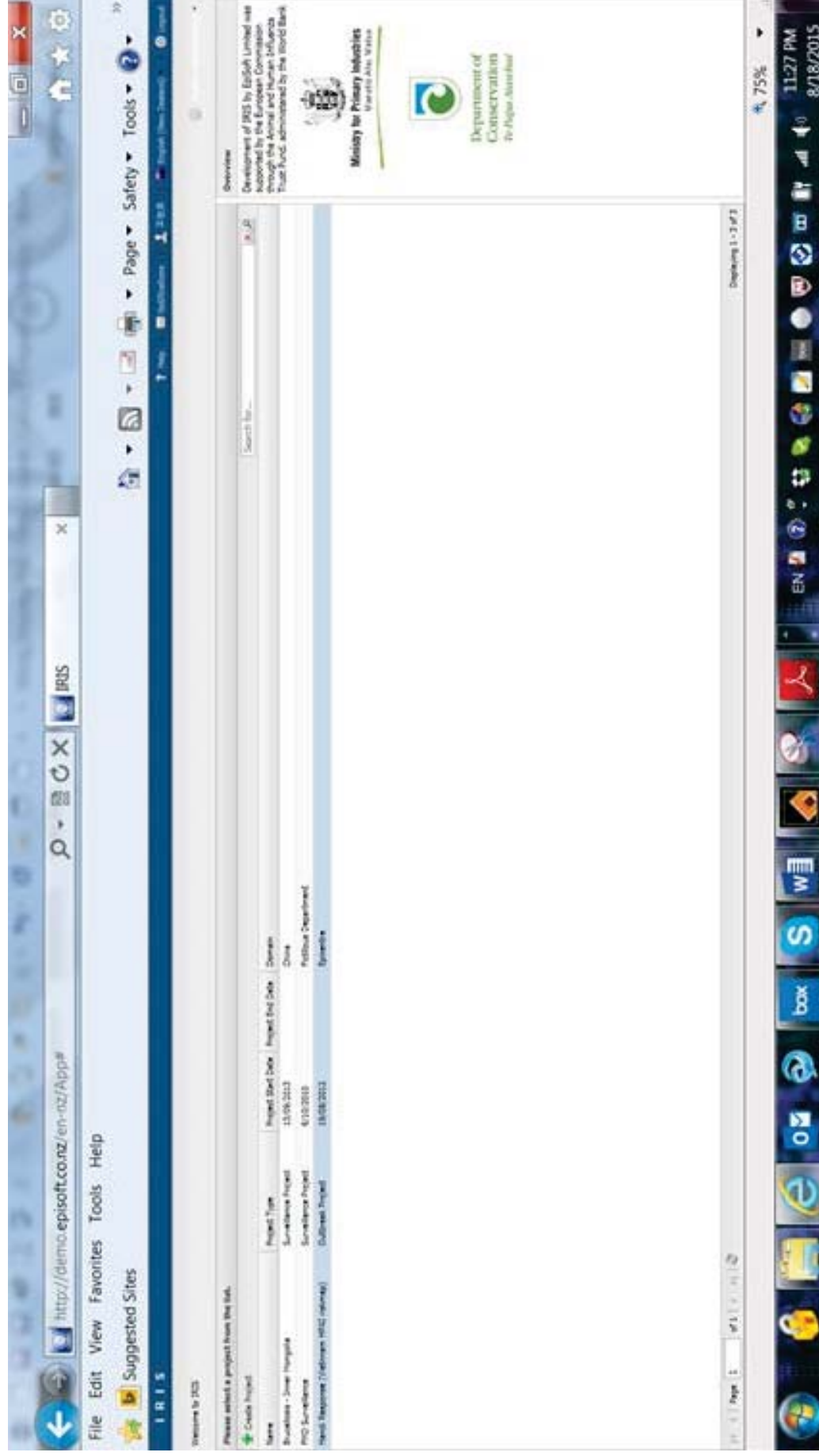
Figure 4-2. Entering HandiResponse

To provide user assistance, an IRIS wiki has been developed. It is available by clicking on "?Help" icon, before entering the term or operation that the user needs to understand.

4.3.2. HandiMap operation

4.3.2.1. Creating or entering into a report

In HandiResponse, the user is presented with all the existing reports by clicking the Tab of "All items". Since the diseases are organized by epitype, the user can also click other tabs/menus such as "Direct Transmission", "Fomite Transmission", etc. so that they will be led to all the existing reports under the selected epitypes. Additional epitypes will need to be added in the next stage of development. A new report can be created by clicking the "add" button, as shown (Figure 4-3).

Figure 4-3. HandiResponse main screen

Procedures to produce outputs from analysis of data stored in a project are called reports. A sub-set of data from a project may be prepared and used for analysis – this procedure is called a filter. As the full range of epitypes is progressively included into the tab options, at least one demonstration (demo) disease risk map report will be provided under each epitype. The demos can be modified by the user, and the report procedure in the demo can be adapted for other diseases which fit under the same epitype category.

The example report shown in Figure 4-4 is the procedure to create a risk map for HPAI H5N1 in southern Vietnam.
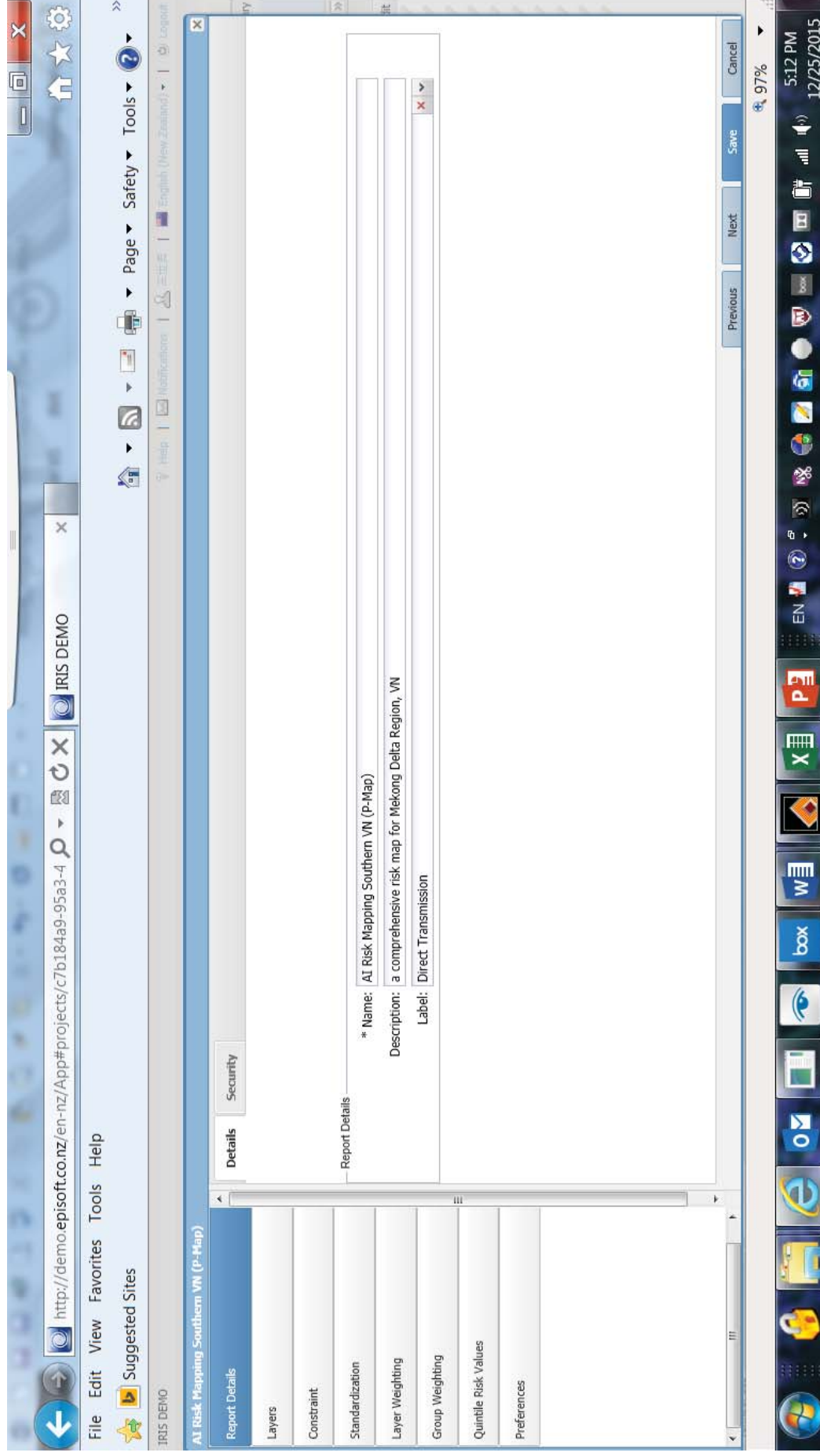
Figure 4-4. Report details on the risk map of HPAI H5N1 outbreak for southern Vietnam

4.3.2.2. Selection of risk layers

At this step, the user is required to choose the risk factors and environmental descriptors/predictors selected for estimating the probability of the disease occurrence by epidemiological unit under study. The risk layers can be either those already available in the HandiResponse data repository, or can be imported from external sources and added to the repository. More than 30 global environmental and climatic layers have already been stored in the repository, and more are being added as the need arises. In additional to global remote-sensed data sets, it is possible to add national spatial data sets.

A zone perimeter needs to be defined for the map boundary before starting selection of risk factor layers. A zone is defined as the geographic area of interest, which could be a country, a subnational administration unit within a country, or a user-defined geographic area covering a number of countries, even an entire continent. In this study, we defined southern Vietnam as an area covering four administrative regions within the country - the South East Region, Mekong River Delta, the South Central Coast Region and the Central Highland Region.

In order to model the disease in HandiSpread on the risk map, the final risk score for each epidemiological unit can be exported and integrated with the "farm file" in HandiSpread, which is the list of epidemiological units, their epidemiological attributes such as animal populations, and their spatial attributes. The risk scores can be exported from HandiMap by clicking "risk index" filter. The file will be created in *.csv* format from the risk map, and can be integrated with the farm file. If HandiSpread finds numerical values in the range zero to 5 in the risk index field of the farm file, it will automatically adjust the susceptibility of each epidemiological unit to establishment of infection, so that high risk units are more likely than the reference level to become infected, while low risk units are less likely.

The risk layers are categorized into groups which have common characteristics. The most frequently used ones such as human population, populations of at risk animal species, transportation/movement, vector density, temperature, altitude, vegetation/landcover types are listed in HandiResponse, while others can be added by the user. Within each group, individual risk layers are selected for inclusion (Figure 4-5). By clicking the magnifier icon, a list of data layers available for choosing from for any specific risk groups will be displayed. A sample list is shown in Figure 4-6.
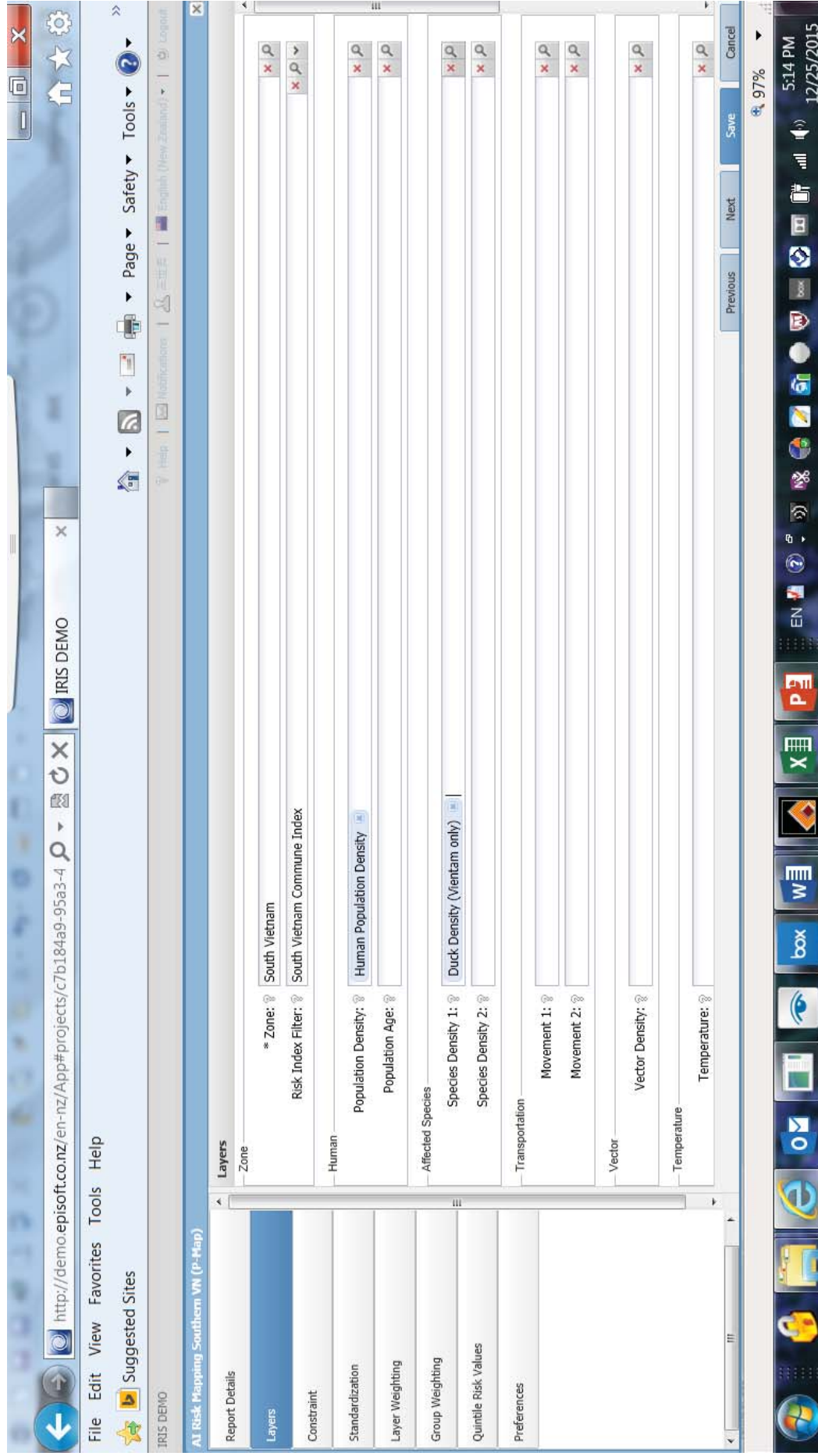
Figure 4-5. Selection of risk groups, layers
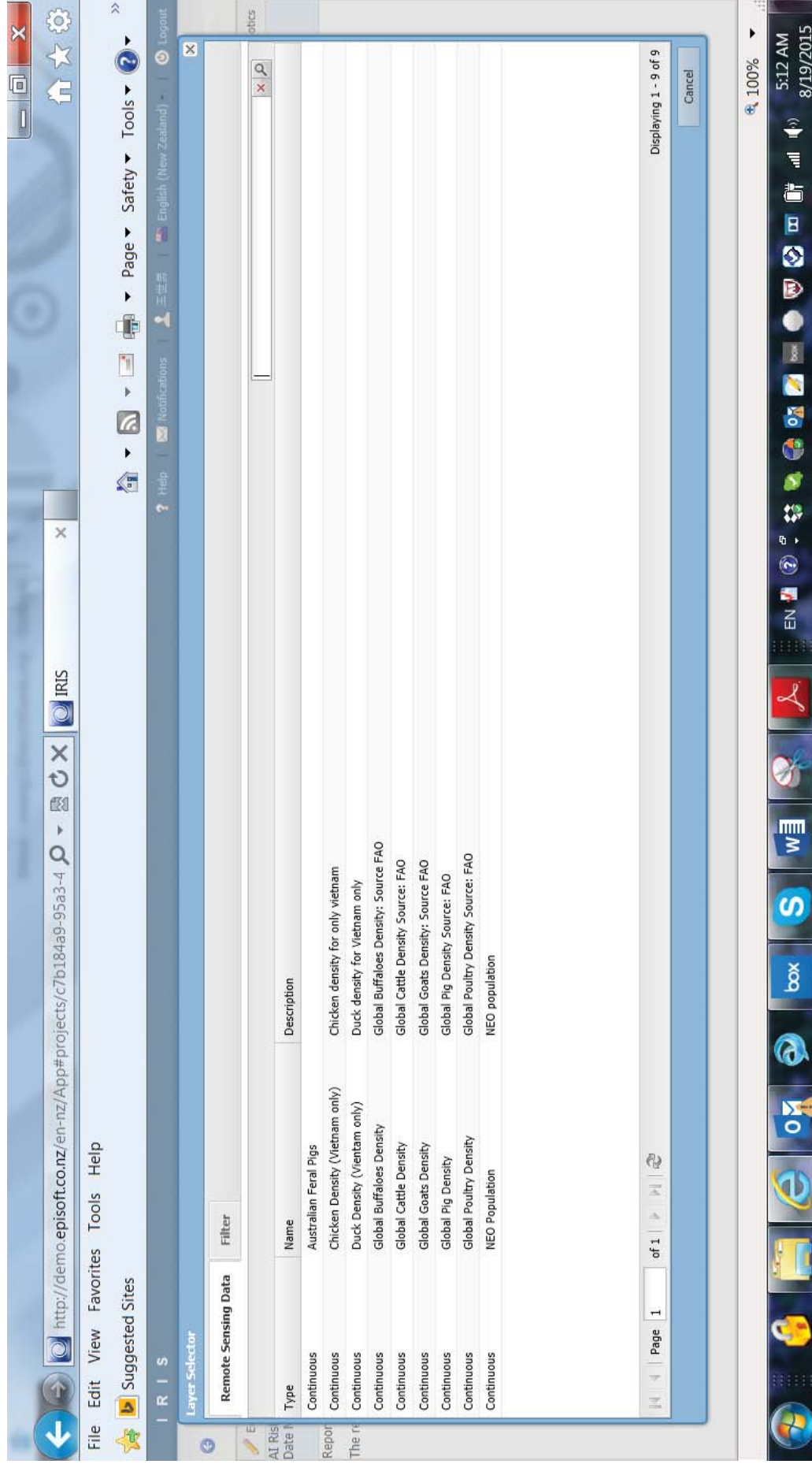
Figure 4-6. List of risk layers for selection under affected species

98

As previously discussed, for risk factor or environmental predictor selection, it is recommended that four principles be applied: (i) epidemiological relevance. Any risk factors or environmental descriptors must be proven or estimated to be correlated with the probability of disease occurrence/presence, either negatively or positively; (ii) the rule of "parsimony". This means a minimum set of inter-independent risk factors and environmental, climatic descriptors should be chosen; (iii) comprehensiveness. The final set of selected risk factors, environmental descriptors should include all the possible ones and (iv) measurability. A numerical value can be assigned to each unit/pixel of a given risk factor and environmental layer.

A thorough literature review is the imperative first step for informing the selection of the most relevant biotic and abiotic risk factors, environment determinants/descriptors. For HPAI H5N1, the following risk layers, environmental descriptors are proposed, based on a literature review (Table 4-1).

Table 4-1. Summary of selected putative risk factors, environmental predictors and their relationship with HPAI H5N1 outbreaks in South East Asia

Risk Factors	Assumed relationship	References
Duck density	Positive, particularly the density of free grazing ducks	Sturm-Ramirez et al. 2005; Hulse-Post et al. 2005; Gilbert et al. 2006; Pfeiffer et al. 2007; Tiensin et al. 2009; Hogerwerf et al. 2010; Gilbert et al. 2011
Chicken density	Increases with chicken density till certain level and then level off, before decreases afterwards	Paul, et al. 2010
Proportion of land used for rice/rice cropping density	Positive association with duck density; Negative association with chicken density	Tiensin et al. 2005; Gilbert et al. 2006; CIVAS, 2006; Pfeiffer et al. 2007; Gilbert et al. 2008; Paul et al. 2010; Hogerwerf et al. 2010; Loth et al. 2011; Gilbert et al. 2011
Distance to highway	Positive association with poultry farming/trading; elevated risk is associated the distance within 10 km to any highways	Gautheir-Clerc et al. 2007; Vannier, 2007; Ward et al. 2008; Paul et al. 2010; Cao et al. 2010; Cao CX et al 2010; Ge et al. 2012
Elevation	Positive or negative association with rice cropping, duck density; low elevation is positively associated with occurrence of HPAI in wild birds	Gilbert et al. 2006; Pfeiffer et al. 2007; Gilbert et al. 2008; Tiensin et al. 2009; Paul et al. 2010, 2011; Si et al. 2010,Takekawa et al. 2010; Loth et al. 2011
Density of human population	Positive (rural pop) or negative (city); medium density is positively associated with elevated HPAI risk	Gilbert et al. 2006, 2008; Tiensin et al. 2009; Loth et al. 2010, Yuniana et al. 2010; Paul et al. 2010; Hogerwerf et al. 2010; Martin et al. 2011

100

In this study, three alternative types of risk maps will be designed by using different risk factor layers and environmental descriptors: environmental descriptors centered map - E-Map, population centered map - P-Map, and movement centered map - M-Map. For E-map, rice paddy coverage, surface water coverage and altitude were selected for HPAI H5N1 risk mapping while for P-Map human population density, duck density and rice paddy coverage were selected. M-Map was the most comprehensive one since almost all risk factors and environment predictors used for P map and E map were included for its development, plus distance to road and distance to water. The selected risk factors and environmental predictors for HPAI H5N1 outbreak in southern Vietnam and their data sources are documented in Table 4-2, and the list of variables used in each map are shown in Table 4-2.

Table 4-2. Selected putative risk factors and environmental predictors for HPAI H5N1 Outbreaks in Vietnam

Risk Factor	Data Source
Human population density	Vietnam Census from MARD
Duck density	Vietnam Census, MARD
Chicken density	Vietnam Census, MARD
Distance to road	GIST, International Steering Committee for Global Mapping (ISCGM) https://gistdata.itos.uga.edu/
Distance to waterway	GIST, International Steering Committee for Global Mapping (ISCGM) https://gistdata.itos.uga.edu/
% of land used for rice cropping	(Xiao et al., 2006; Xiao et al., 2005)
% of land covered by surface water	GADM at http://www.gadm.org
Elevation	Data available from the U.S. Geological Survey. http://topotools.cr.usgs.gov/GMTED_Viewer/

Note: MARD, Ministry of Agriculture and Rural Development, Viet Nam.

4.3.2.3. Setting Constraint(s)

Constraints can be added for each risk layer. For instance, certain environments may be excluded from the risk map - such as mountain regions which are snow covered all year round, locations above a certain altitude or latitude, desert, etc., that are considered unsuitable for occurrence of the particular disease. They need to be screened out at this stage. For HPAI H5N1 in Southern Vietnam, no constraints were set, because all areas were considered suitable for occurrence of the disease.

4.3.2.4. Standardization of value of each risk layer

Standardization has two purposes: firstly, the minimum and maximum of raw values per risk layer varies. The raw values need to be converted to a standard scale so that values are comparable before any logic operations could be performed to generate a risk map. In HandiRepsonse, a scale can be chosen to suit the situation - from 0 to 1 or to10 or to 100, etc. for standardization, as long as all layer variables are converted to the same scale. For the HPAI H5N1 risk map, the range was chosen to be

between 0 and 1. This means all the values of pixels for each risk layer have to be rescaled to take values between 0 and 1, with 0 representing no risk and 1 representing the maximum risk. Secondly, a relationship needs to be assumed between the scale in each layer and the risk or probability of disease occurrence. This may be linear or non-linear, and may be direct or inverse. Figure 4-7 demonstrates the result for the duck density layer standardization.
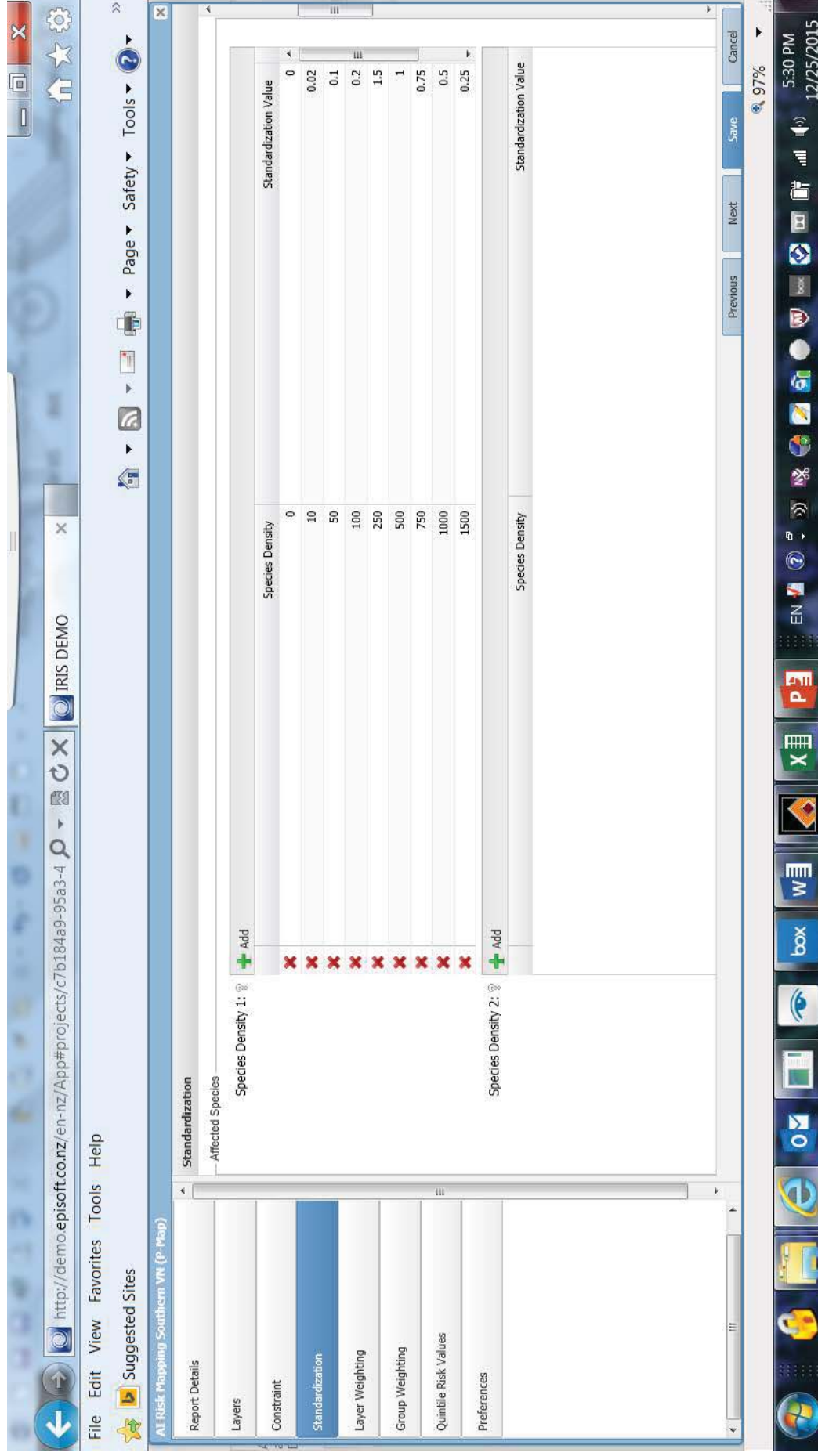
Figure 4-7. Results of standardization of duck density layer

Not all factors exert the same intensity of influence on occurrence of the disease, so after standardization a weight must be assigned to each factor according to its perceived importance in influencing disease occurrence. This is a very important part of the process, which relies on evidence from the scientific literature and other sources. Weight assignment is designed to be accomplished in two steps in HandiMap: step one is to assign weight for each risk factor/environmental descriptor within the given risk group; step two is to assign a weight for each risk group in relation to other risk groups. The advantage of the two step weight assignment is that it is effective in handling a large number of risk layers, and it is essential for factors such as vegetation type where there are multiple layers within a group, but they are mutually exclusive for each pixel or raster unit. The risk layer and group weight assignments for M-Map is illustrated in Figure 4-8 and Figure 4-9 respectively. Weights within a block of factors must sum to 1.0. Analytic Hierarchy Process (AHP) is recommended to be used to facilitate the weight elicitation.   The AHP approach has been a widely used weight elicitation method and there exist ready-made programs to carry it out (Ossadnik & Lange, 1999). A sample weight elicitation using the analytical hierarchy process (AHP) in an Excel spreadsheet is documented in Table 4-3.

Table 4-3. AHP for weight assignment for risk layers for the P-Map estimating the risk of HPAI H5N1 outbreak, southern Vietnam

	A	B	C	D	F
1	**Risk factor/proxy**	**Movement**	**Environment**	**Affected Species**	
2	Movement	1	2	7	
3	Environment	0.50	1	4	
4	Affected Species	0.14	0.25	1	
6					
8	**Risk factor/proxy**	**Movement**	**Environment**	**Affected Species**	**Weight**
9	Movement	0.61	0.62	0.58	0.60
10	Environment	0.30	0.31	0.33	0.32
11	Affected species	0.09	0.08	0.08	0.08
12					
13	**Lambda max**	3.003			
14	**CI**	0.001			
15	**CR**	0.002	<10%		

Note: CI stands for consistency index; CR stands for consistency ratio. It is a comparison between consistency index and random consistency index. In case CR is equal or less than 0.1, the inconsistency level is acceptable If CR is less than 10%, the level of inconsistency is acceptable.
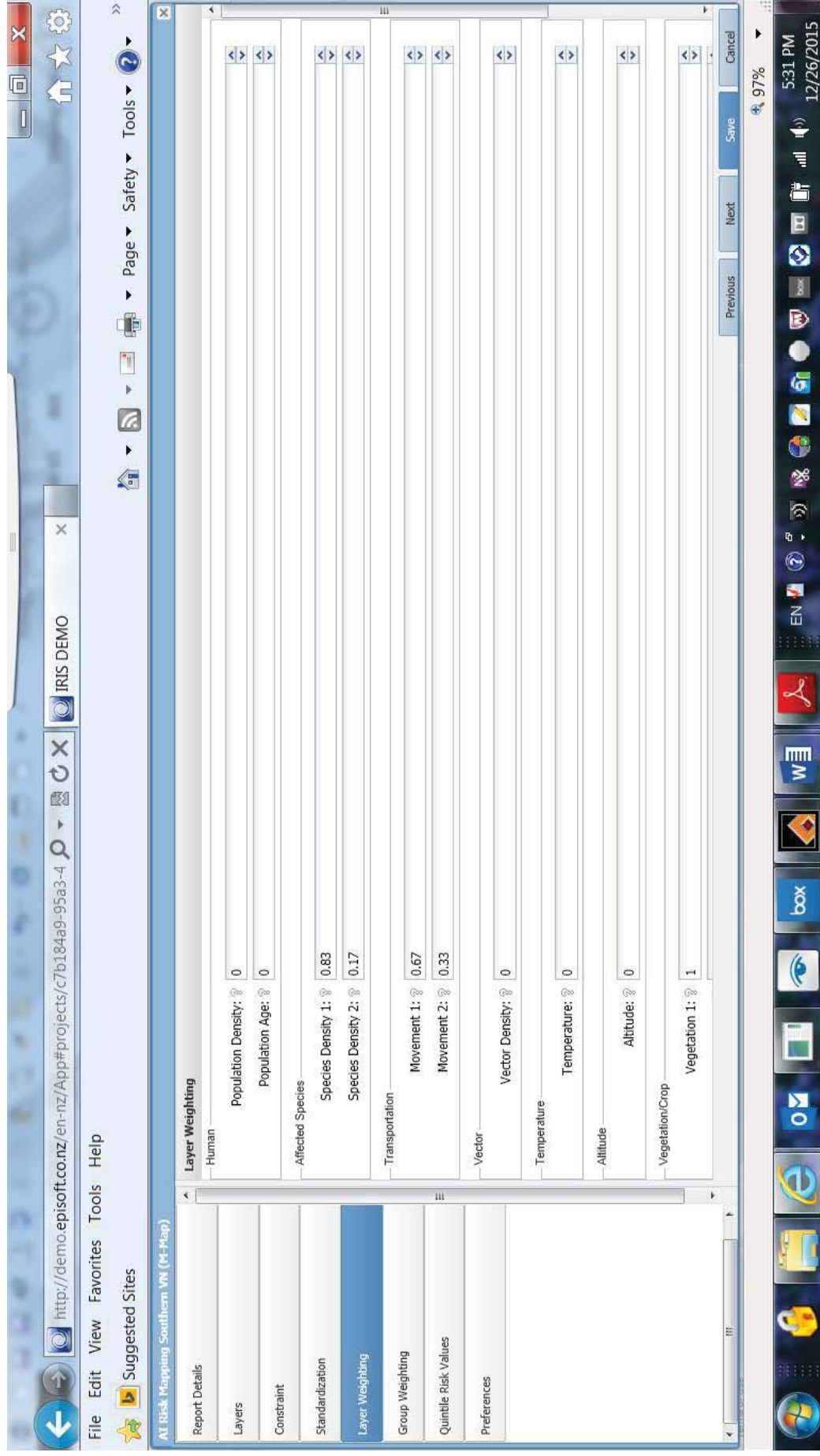
Figure 4-8. Risk layer weight assignments of M-Map for HPAI H5N1 outbreak, southern Vietnam
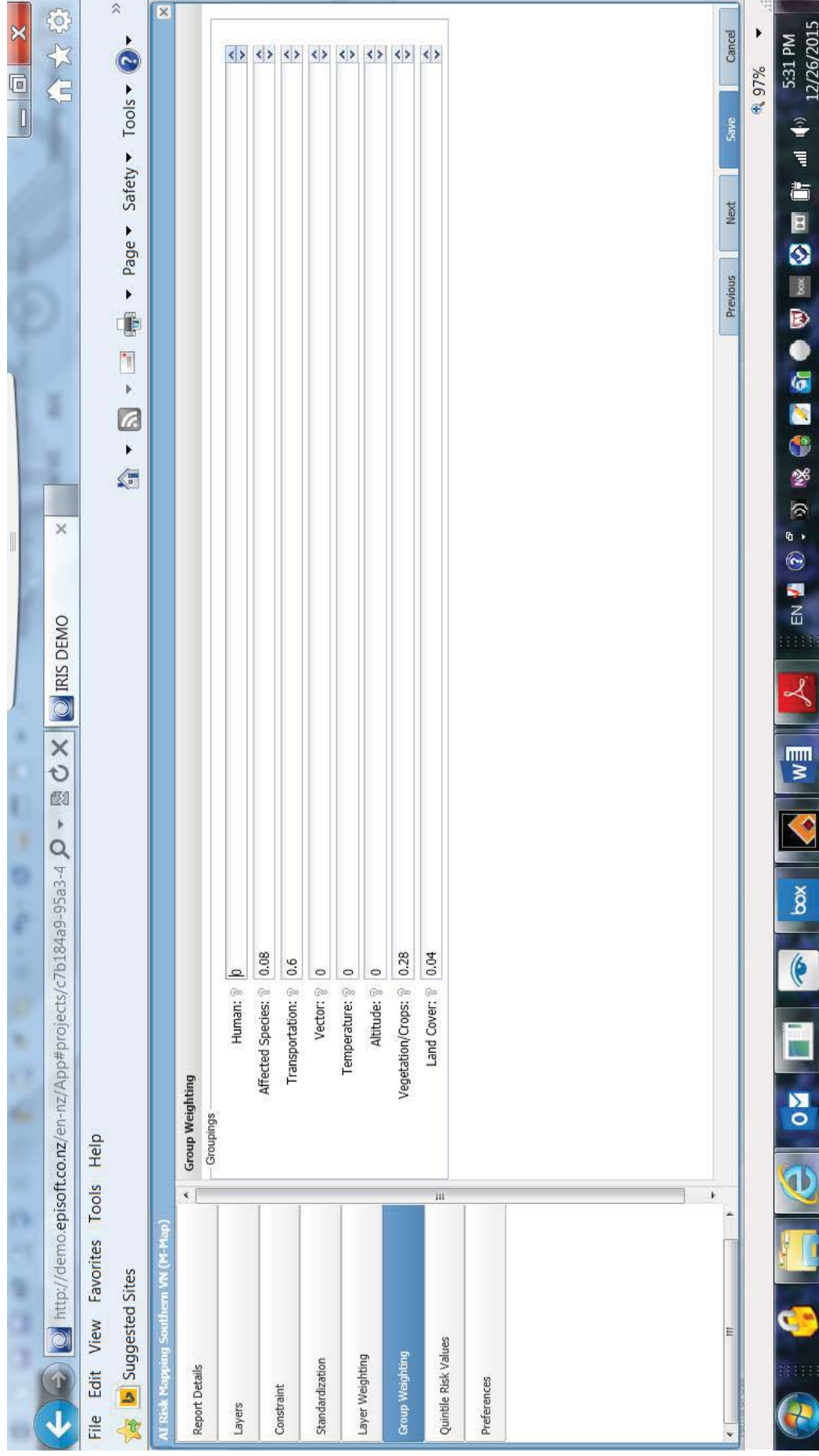
Figure 4-9. Group weight assignments of M-Map for HPAI H5N1 outbreak, southern Vietnam

Weight assignments for all the selected risk factors and environmental predictors used for three maps are summarized in Table 4-4.

Table 4-4. Final weight assignments for selected risk factors and environmental descriptors

Risk factor	E-Map	P-Map	M-Map
Human population density		0.10	
Duck density		0.09	0.07
Chicken density			0.01
Distance to highway			0.40
Distance to waterway			0.20
% of land covered by rice paddy	0.80	0.81	0.28
% land covered by water	0.11		0.04
Altitude	0.09		
Consistency ratio	8.99E-04	9.76E-05	0.002

Notes: Consistency ratio (CR) is a comparison between consistency index and random consistency index. In case CR is equal or less than 0.1, the inconsistency level is acceptable.

4.3.2.6. Risk categorization of epidemiological units

In principle risk can be assigned at the smallest raster unit available, which may be as small as a pixel, but in the case of satellite data is typically a square km, or 5 sq km, etc. However typically it is necessary to assign risk values to epidemiologically and managerially meaningful spatial units, such as farms, villages, or in the case of HPAI in Vietnam, communes, which are groups of 10 to 15 villages. These are vector rather than raster units, and can be processed either as a shape file or as a point location situated within the spatial unit, which represents the unit as a whole. This point or the centroid of a polygon is considered to represent the risk level of that epidemiological unit.

In HandiMap, all epidemiological units within the map zone are divided into quintiles according to their risk levels. Transmission probabilities (effectively the susceptibility of a spatial unit to becoming infected) are then adjusted from the reference level so that high risk spatial units are more likely than the reference level to become infected (and hence transmit infection onward),while low risk units are less likely. Users can choose any of the five quintiles as the reference level and assign it a risk adjustment value of 1, but the 3rd quintile would usually be appropriate, or the 2nd quintile offers an alternative, as shown below. The rest of the quintiles are then assigned a risk adjustment value above or below the reference level. The adjustment factor may be linear, or non-linear, typically with risk escalating upward from the reference level to level 5, but linear downward to level 1 (termed a rapid increase scheme). The results of risk adjustment values for all epidemiological units can then be exported as a *csv* file to be used as an input for modelling spatial distribution of a disease outbreak, or to guide the development of any risk-based surveillance strategies.

A "5 Percent Rapid Increase" scheme for risk level categorization is illustrated in Figure 4-10. For a HPAI H5N1 risk map for southern Vietnam, the 3rd quintile is set to be the reference level with a risk adjustment value of 1. Under this particular rapid increase scheme, the risk adjustment value for the 4th quintile is set to 1.05 (or 1+0.05) and the fifth quintile to 1.15 (or 1.05+0.05×2). For each quintile below the 3rd quintile, proportional or linear decrease is applied. The risk adjustment value for the 2nd Quintile is 0.95 (or 1-0.05) and that for 1st quintile 0.9 (or 0.95-0.05). The choice of risk adjustment scheme and the scaling factors used for adjustment are both a matter of epidemiological judgment, and the effect of changing them can easily be explored, as shown in Chapter 5.
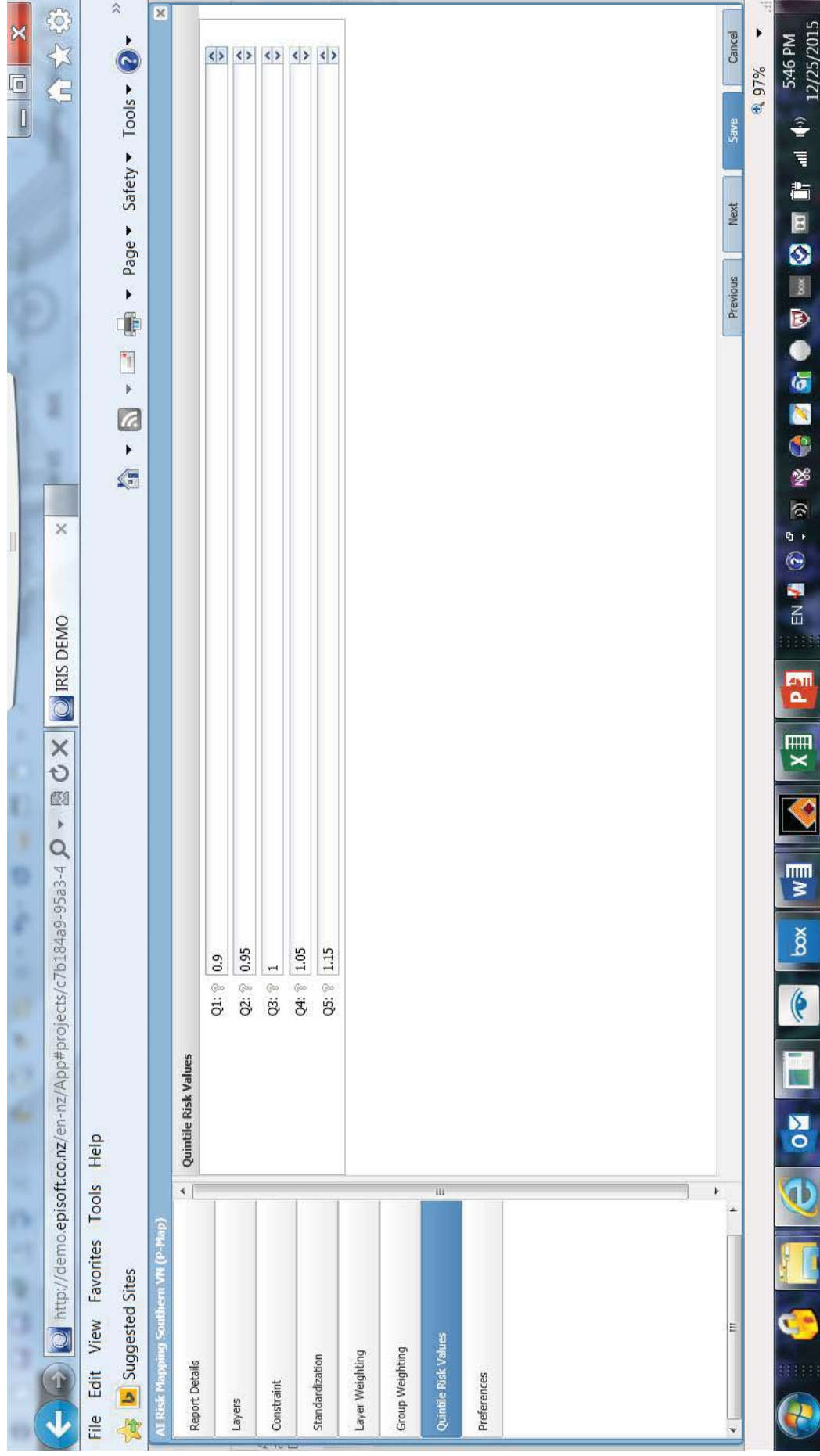
Figure 4-10. A sample of risk level categorization

4.3.2.7. Selection of map display type and color scheme for map presentation

In this step, the user can choose how the disease risk map will be presented. By default, HandiResponse assumes a kernel smoothed map is the preferred choice. The underlying map can be any of the four choices - Open Street, Google Street, Google Hybrid and Google Satellite. There are 14 color schemes available for the users to choose from for map presentation (Figure 4-11), to achieve different visual representations of risk levels.
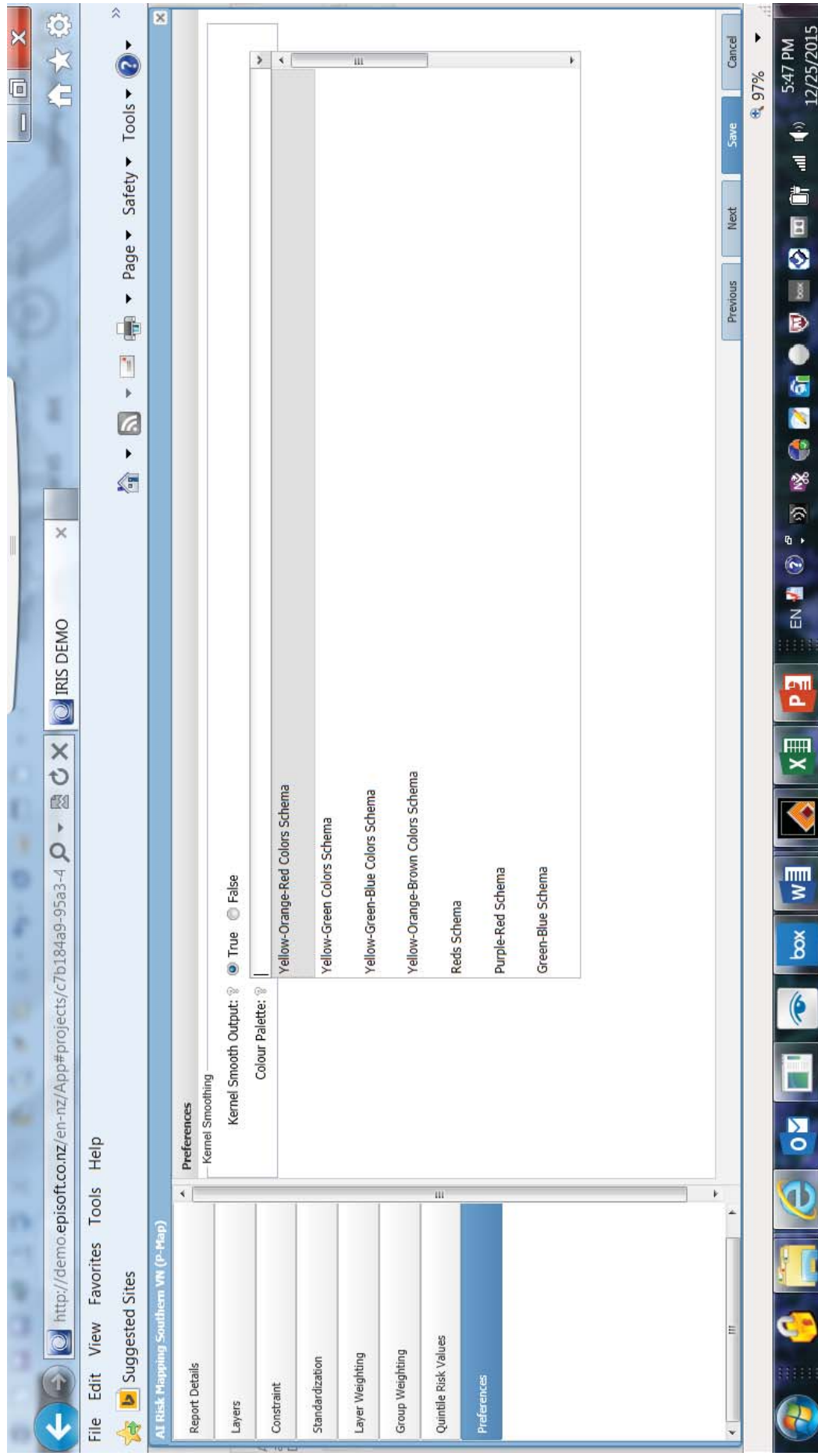
Figure 4-11. Selection of map presentation

111

This completes all the steps to create a disease risk map. The last but the most important step is to click "save". HandiResponse will then do the all necessary calculations and processes, then will generate a risk map. This takes a few seconds.

4.4. Results

Maps produced by HandiResponse can be either kernel smoothed (Figure 4-12), or raw pixel (non-kernel smoothed) versions (Figure 4-13); with the underlying landscape being presented as one of four user-selected choices - open street, google street, google hybrid and google satellite (Figure 4-14). Apart from integrated risk maps, the user can also choose to display any individual risk layers (Figure 4-15).

Based upon different individuals' judgments about the epidemiological importance of particular risk factors and environmental descriptors, different risk maps can be developed and the importance of the differences evaluated. In the case of the HPAI H5N1 risk maps in southern Vietnam, three risk maps were developed: P-Map (Figure 4-12, Figure 4-13 and Figure 4-14), E-Map (Figure 4-16) and M-Map (Figure 4-17). They represent potential different findings and viewpoints on the importance of the various risk factors, as reported in the scientific literature.
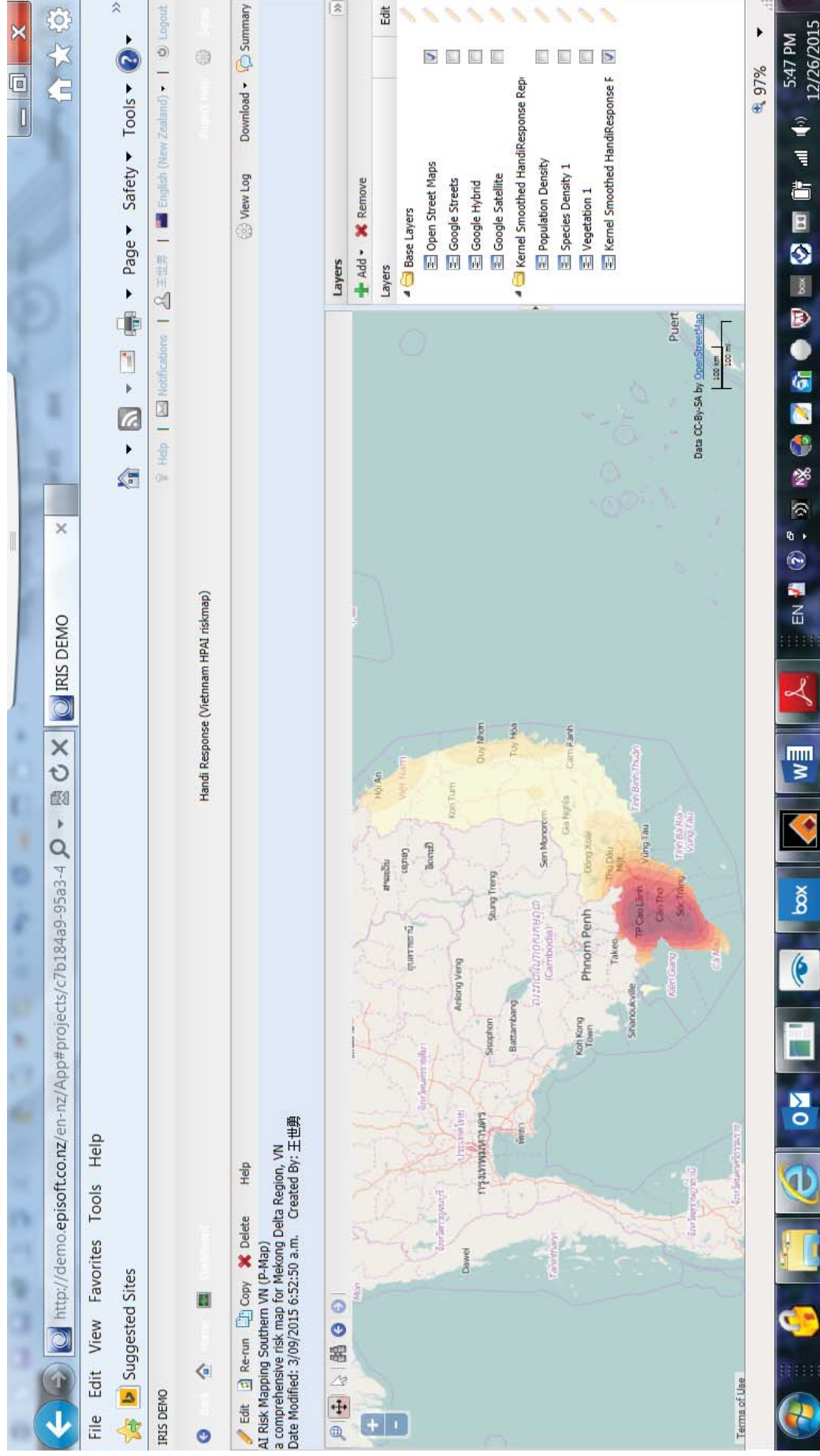
Figure 4-12. Kernel smoothed P-Map for HPAI H5N1 outbreak, southern Vietnam (open street map)

Note: The risk map is based on duck density, rice paddy coverage and human population density risk layers; the darker the color, the higher the risk for HPAI H5N1 outbreak.

113
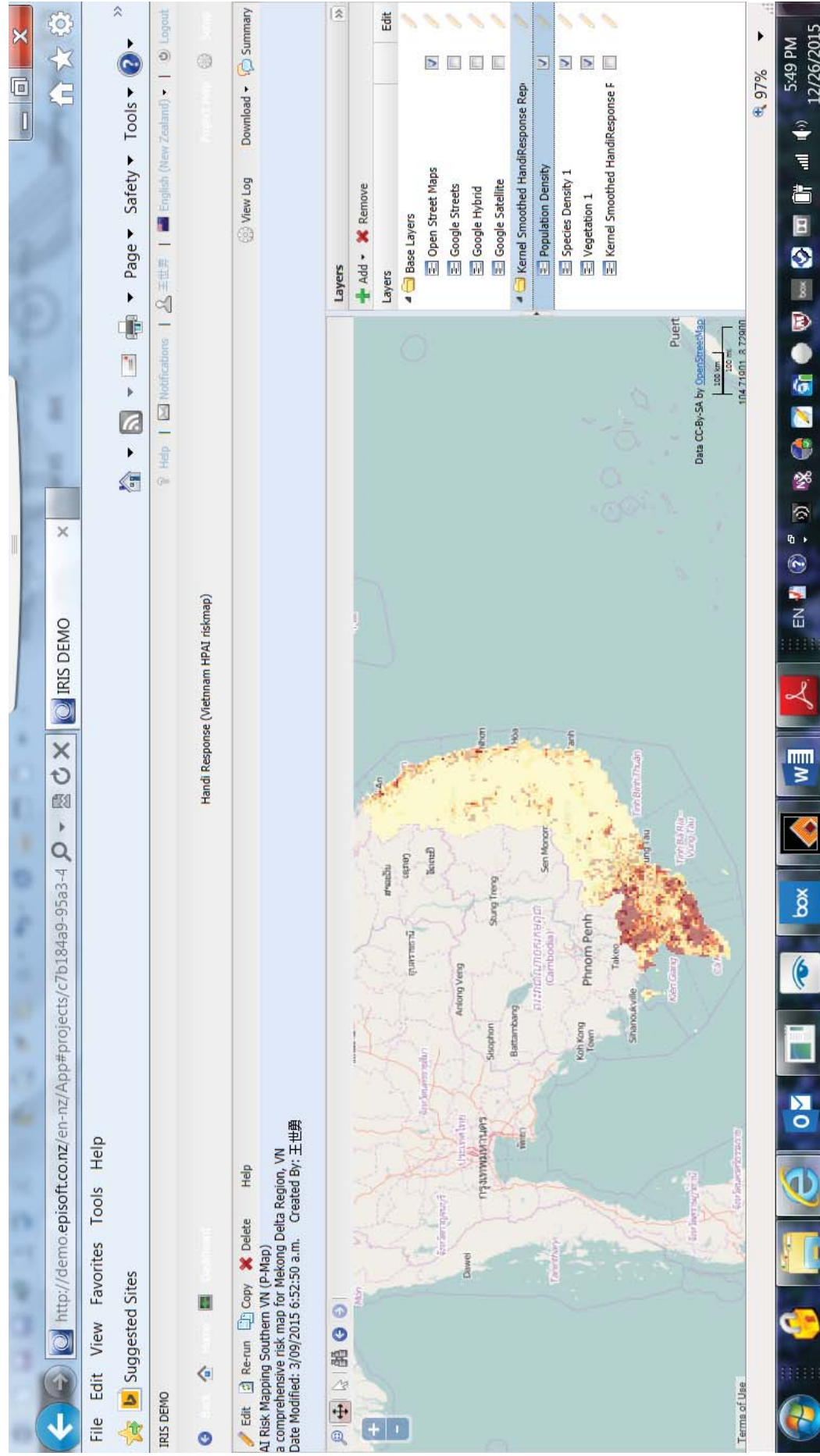
Figure 4-13. Non-kernel smoothed P-Map for HPAI H5N1 outbreak, Southern Vietnam (open street map)
Note: the darker the color, the higher the risk for HPAI H5N1 outbreak.

114

A: P Map (Google street)   **B: P Map (Google hybrid)**   **C: P Map (Google satellite)**
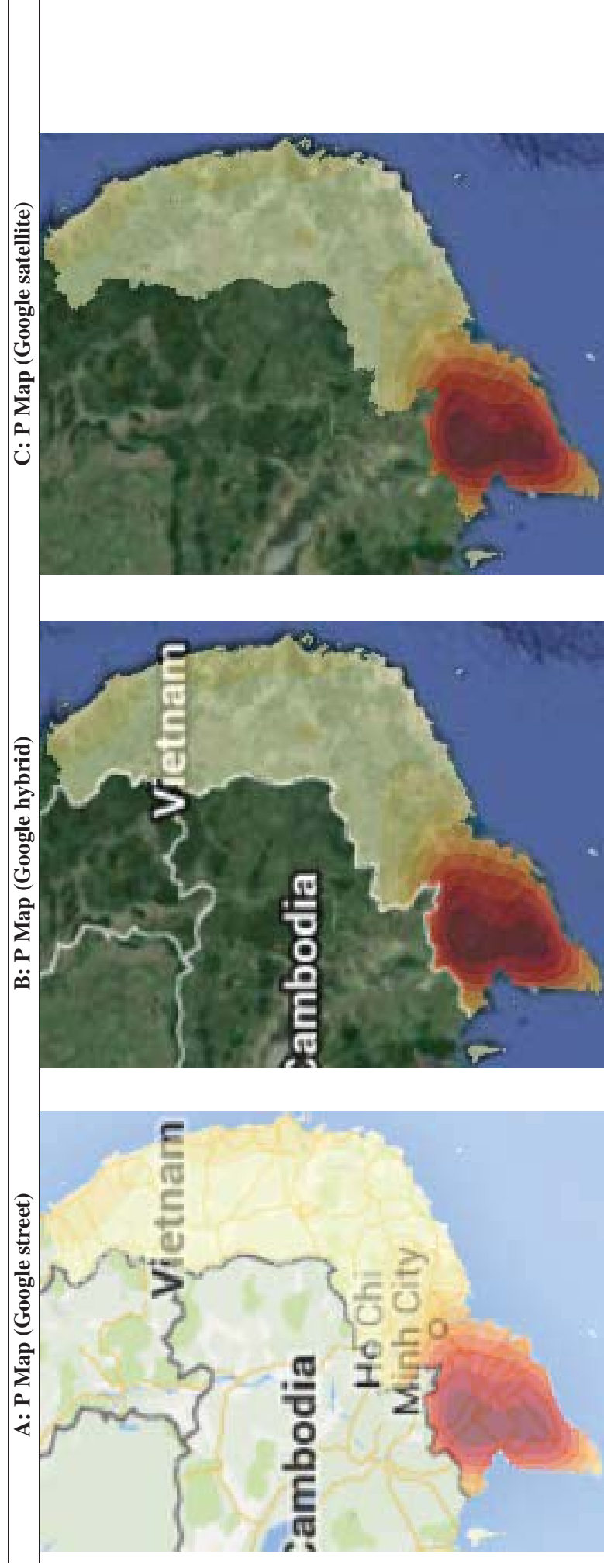
Figure 4-14. Options for map presentation by HandiResponse
Note: the darker the color, the higher the risk for HPAI H5N1 outbreak.

A: Rice paddy coverage

B: Human density

C: Duck density
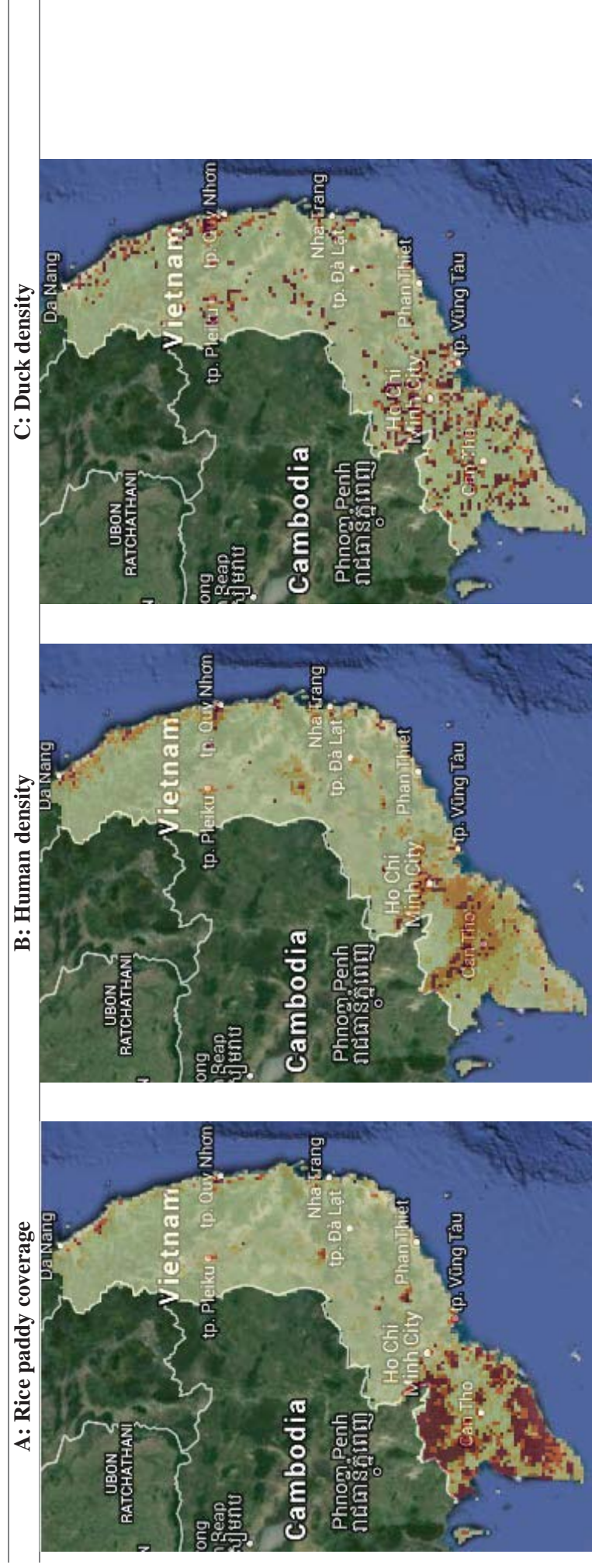
Figure 4-15. Selected individual risk layers for HPAI H5N1 outbreak in southern Vietnam

Note: the darker the color, the higher the risk for HPAI H5N1 outbreak.

116

A: Smoothed E Map   B: Unsmoothed E Map
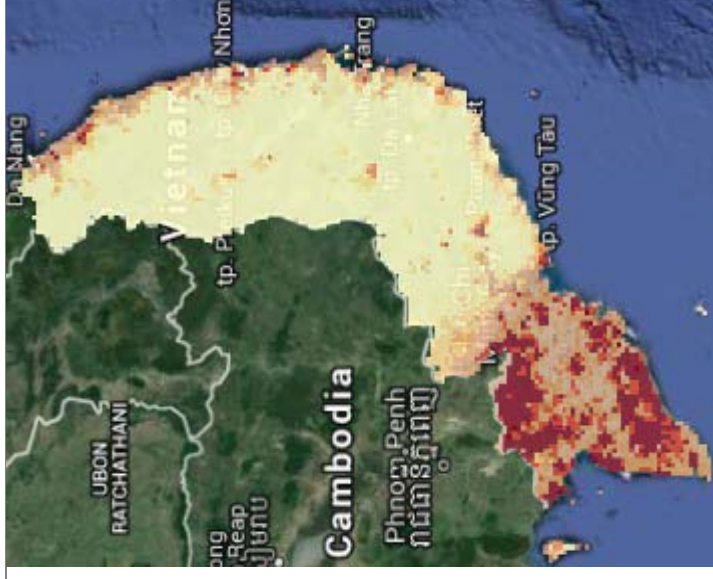
Figure 4-16. E-Map for HPAI H5N1 outbreak in southern Vietnam
Note: the darker the color, the higher the risk for HPAI H5N1 outbreak.

117

A: Smoothed M Map

B: Unsmoothed M Map
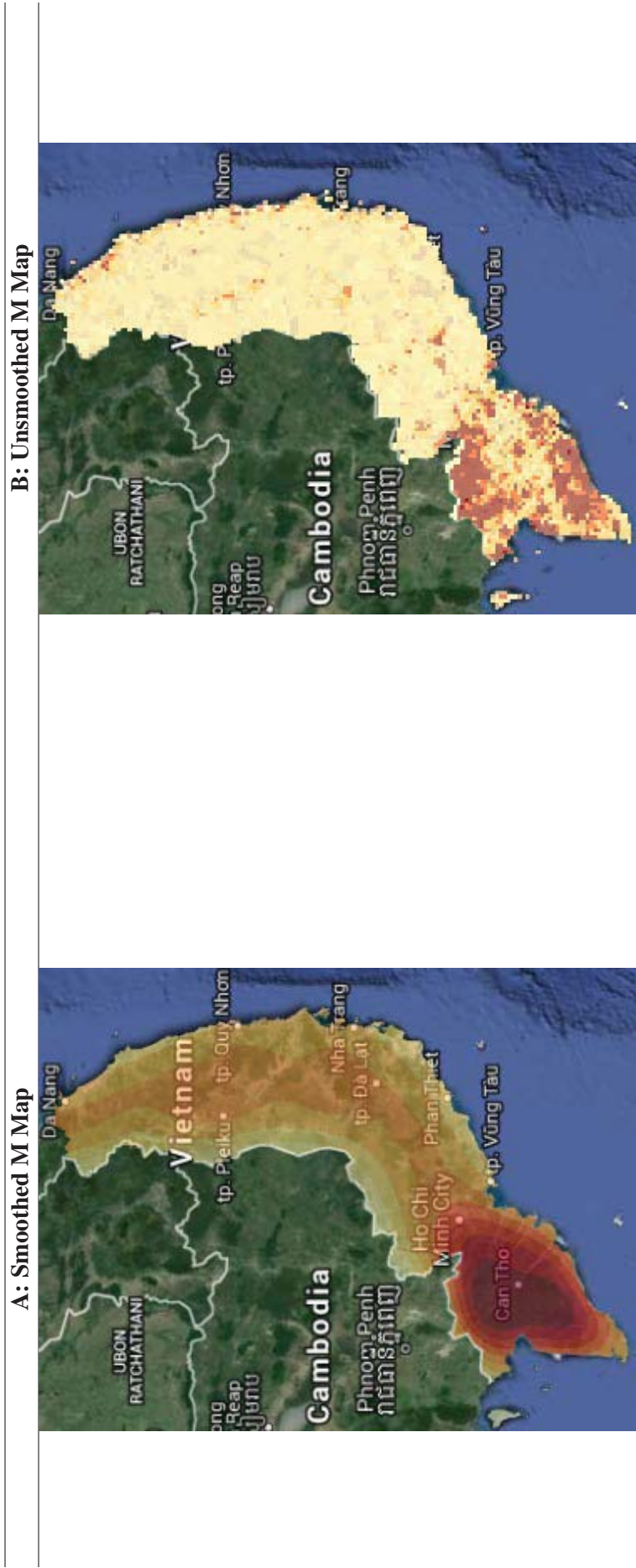
Figure 4-17. M-Map for HPAI H5N1 outbreak in southern Vietnam
Note: the darker the color, the higher the risk for HPAI H5N1 outbreak.

118

As mentioned earlier, HandiMap can also export the risk adjustment values of the risk landscape to a csv file. The risk file has four columns: ID number of each epidemiological unit, corresponding risk score, as well as its northing and easting (Figure 4-18).

	A	B	C	D
1	ID	Risk	X	Y
2	7150503	0	108.4422	11.43594
3	6010303	0	107.6501	15.30661
4	7050411	0	108.7369	11.92642
5	6052525	0	108.5749	12.37282
6	6071305	0	107.5965	11.73409
7	6070508	0	108.4234	12.21869
8	5111115	0	108.7483	12.24837
9	5031305	0	107.5275	15.7175
10	5090303	0	108.8537	13.46978
11	7050327	0	108.6581	11.60045
12	5030407	0	107.3663	15.80327
13	6010317	0.001	107.5744	14.96488
14	6051725	0.001	108.7754	12.54399
15	6031303	0.001	108.41	13.89064
16	5032827	0.001	108.0751	15.17623

Figure 4-18. The structure and content of a sample risk file produced by Handimap
Note: ID stands for commune identifier; Risk stand for HPAI H5N1 risk; X stands for X coordinate and Y stands for Y coordinate.

4.5. Discussion

4.5.1. Rationale for development of HandiMap

As a valuable technique in spatial epidemiology, disease risk mapping can help test a hypothesis, identify gaps in our knowledge, provide a direction for surveillance and control efforts, or evaluate the actual or potential effectiveness of an intervention (Brownstein, Freifeld, & Madoff, 2009; Kitron, 2000; Tatem, Smith et al., 2010). It can also provide a valuable evidence base for achieving progress towards global health commitments (Pigott, Howes et al., 2015). However, a recent review revealed that out of 176 globally important infectious diseases with a strong rationale for mapping, only 4% have been adequately mapped (Hay, Battle et al., 2013).

Compared with other commercially available mapping tools, HandiMap has the following merits: (i) at least one disease risk map under each epitype has been developed or will be developed for program users to learn by "mirroring" from the demonstration disease to the disease of interest to them; (ii) more than 30 global environmental and climatic descriptor layers have been stored in the IRIS data server and can be accessed freely, so an epidemiologist can explore the use of different layers without needing to download

119

them from multiple servers and install them in a standard GIS package. The number of layers will continue to increase, and the file sizes are very large so it is much better to have them stored on a single site than to have individual users store them on local hard drives. The layers will also need to be regularly updated as new releases become available, and this is best done through a single server site; (iii) risk values by epidemiological unit can be exported and used by other modules of HandiResponse (or other disease modelling programs) to simulate disease spread, and assess the effect of any disease surveillance and mitigation strategies and (iv) the tool is free of charge.

4.5.2. Selection of risk layers and environmental descriptors for genuinely novel diseases

If the disease of concern involves an unknown pathogen for which only limited evidence is available, it is critical to collect and analyze the epidemiological information on the disease through outbreak investigation, sporadic case investigation, intervention studies and expert elicitation approaches (Pires, Evers et al., 2009). Besides, ex ante risk assessment on possible emerging zoonotic diseases and prioritization exercise is highly recommended for any countries that wish to improve their preparedness for these diseases (Balabanova, Gilsdorf et al., 2011; Cardoen, Van Huffel et al., 2009; Doherty, 2000; Andreas Gilsdorf & Gérard Krause, 2011; Havelaar, Van Rosse et al., 2010; Krause, 2008; Kurowicka, Bucura et al., 2010; Mangen, Batz et al., 2010; McKenzie, Simpson et al., 2007). This process will help to identify the top ranking emerging infectious diseases relevant to any given countries, so that information collection on them can be implemented well in advance.

The following practical questions need to be considered before selection of risk factors for a given disease risk mapping: (i) whether the disease affects any specific sub-group(s) of the human population; (ii) what species of animals are affected; (iii) whether the disease affect specific sub-groups within animal populations; (iv) whether any vectors are involved; (v) whether any products or specific foods are implicated in disease spread; (vi) whether the disease is explicitly or implicitly trade-related; (vii) whether the disease incidence varies seasonally and (viii) whether the disease occurrence or spread is influenced by any particular environmental factors.

In this study, we developed three types of disease maps which make use of particular combinations of putative risk factors, namely E-Map, P-Map and M-Map.

In the E-Map, we selected percentage of land covered by rice paddy, altitude and percentage of land covered by water to predict the risk of HPAI H5N1 outbreak (Desvaux, Grosbois et al., 2011; Ge, Haining et al., 2012; Gilbert, Chaitaweesup et al., 2006; Gilbert, Xiao et al., 2008; Iglesias, Muñoz et al.,2010; Loth, Gilbert et al., 2011; Martin, Pfeiffer, et al., 2011; Paul, Tavornpanich et al., 2010; Paul, Wongnarkpet et al., 2011; Pfeiffer, Minh et al., 2007; Si, Wang et al., 2010; Takekawa, Newman et al.,

2010; Tiensin, Ahmed et al., 2009; Van Boeckel, Thanapongtharm, et al., 2012; Yupiana, de Vlas et al., 2010). All these environmental descriptor layers were extracted from the remote-sensed images. They represent a practical solution for estimating HPAI H5N1 outbreak risk in case (i) there is no readily available agricultural census data; (ii) the data is available but out of date and (iii) data are only available with coarse spatial and temporal resolutions (Gilbert, Xiao et al., 2007). Another advantage of the E-model is that data obtained have less chance to be manipulated. E-model was grounded on the understanding that rice cropping intensity was one of the most important predictors for the HPAI H5N1 outbreak in southern Vietnam, by influencing both duck density and duck movement patterns. Gilbert et al. claimed that cropping intensity was a better epidemiological descriptor for the density of free-grazing ducks than the reported duck density, since the latter was obtained from a duck registry done at the home villages, but ducks may migrate to other villages with rice paddy and harvests where they could get infected and spread the disease (Gilbert, Xiao et al., 2008; Xiao, Gilbert et al., 2007). We used percentage of land covered by rice paddy to represent cropping activities since we could not get the data on rice cropping intensity. Besides, human population density was used as a surrogate or an explanatory variable for anthropogenic factor(s) such as poultry movement in this map (Gilbert, Xiao et al., 2008).

In P-Map, as well as selection of rice paddy as an environmental predictor, other risk factors such as duck density and human population density were introduced. Human density was used as a surrogate representing disease risk related to human behaviours affecting poultry rearing and movement and also the influence of closeness of interaction between families and their poultry flocks (Gilbert, Chaitaweesup et al., 2006; Gilbert, Newman et al., 2010; Gilbert & U Pfeiffer, 2012; Gilbert, Xiao et al., 2008; Hogerwerf, Wallace et al., 2010; Loth, Gilbert et al., 2011; Paul, Tavornpanich et al., 2010; Pfeiffer, Minh et al., 2007; Sturm-Ramirez, Hulse-Post et al., 2005; Tiensin, Chaitaweesup et al., 2005).

M-map employed almost all risk factors and environment predictors used for P map and E map so that it could be considered to cumulate the contents of the two previous risk maps (chicken population density, duck population density, percentage of commune covered by rice paddy, percentage of commune covered by surface water) and add two surrogate variables for mobility of people and their poultry (inverse of distances to road and waterway). The selection of these factors was based on reviewing relevant literature (Cao, Xu et al., 2010; Fang, de Vlas et al., 2008; Gilbert, Xiao et al., 2008; Martin, Pfeiffer, et al., 2011; Peterson & Williams, 2008; Pfeiffer, Minh et al., 2007; Rivas, Chowell et al., 2010; Tiensin, Chaitaweesup et al., 2005; Williams & Peterson, 2009; Xiao, Boles et al., 2005; Yupiana, de Vlas et al., 2010). Poultry movement was confirmed to be significantly associated with the spread of HPAI H5N1 outbreaks in Vietnam (Pfeiffer, Minh et al., 2007). The two surrogate variables for movement were considered as the most important risk factors in the M-Map, based upon this evidence and others from a

wider literature review (Cao, Xu et al., 2010; Gauthier-Clerc, Lebarbenchon, & Thomas, 2007; Ge, Haining et al., 2012; Gilbert & Pfeiffer, 2012; Paul, Tavornpanich et al., 2010; Pfeiffer, Minh et al., 2007). While they are quite indirect ways of measuring mobility, they are the most practical to use when no direct movement data is available. Rice cropping was again included and considered just as potentially influential as the movement risk factors. Duck density, chicken density, percentage of land covered by surface water were also enlisted as the risk factors or environmental predictors. Percentage of land covered by surface water was proven to be correlated with elevated risk for HPAI H5N1 due to a putative association with presence of wild birds, density of duck population or contact between domestic water birds and wild birds (Desvaux Grosbois, et al., 2011; Martin, Pfeiffer, et al., 2011; Van Boeckel, Thanapongtharm et al., 2012; Van Boeckel, Thanapongtharm et al., 2012).

Other environmental and climatological predictors such as precipitation, normalized difference vegetation index (NDVI) had also been confirmed as having an association with HPAI H5N1 outbreaks in different locations (Fang, de Vlas et al., 2008; Henning, Henning et al., 2009; Peterson & Williams, 2008; Si, Wang et al., 2010; Williams & Peterson, 2009). Henning et al. used re-scaled NDVI value of 140 to 160 to represent land consisting of waterway with widespread vegetation. They hypothesized that such a geographic landscape was associated with ducks and an abundance of a variety of wild birds associated with these wetlands that might pose a further risk for HPAI occurrence (Henning & Pfeiffer, 2009). They concluded that in Vietnam, medium level May to October NDVI was associated with higher HPAI risk. By using ecological niche modelling, Williams described that HPAI H5N1 was predicted to be absent from areas with low NDVI values and low seasonal variation, but present in areas with marked seasonal variation in their Middle East and northeastern Africa assessment, with an exception for the Arabian peninsula (Williams & Peterson, 2009). These two studies indicate that using NDVI to predict HPAI H5N1 is highly locality specific. The key is to identify what underlying epidemiological factors NDVI is intended to represent. Fang et al. identified that annual precipitation was negatively associated with HPAI H5N1 outbreaks in China. They hypothesized that lower precipitation levels may lead to a higher concentration of birds in a reduced number of wetlands, thus increasing the chances of birds becoming infected through contact with the virus (Fang, de Vlas et al., 2008).

Following the rule of parsimony on selection of the layers for risk map development, we decided not to use NDVI since we were able to use data on rice paddy and surface water as more directly relevant risk factors, and besides, NDVI is a very short term measurement. We also did not use annual precipitation data as because of the short duration of the epidemic waves, clearly other factors were more influential than this.

122

4.5.3. Standardization of risk values

HandiResponse does not specify any particular standardization methods to be used. It actually accommodates various approaches for standardization, such as linear scale transformation methods like maximum standardization, interval standardization and goal standardization and non−linear value function approaches (Drobne & Lisec, 2009; Eastman, 1997; Lai & Hopkins, 1989; Malczewski, 1999; Voogd, 1983).

For a meaningful standardization of various risk data, it is desirable to understand the distribution of the values of each risk factor, the minimum, median and maximum value. Secondly, standardization also means that a relationship or function between raw metrics and disease occurrence/presence needs to be assumed. Establishment or estimation of a relationship must be based upon comprehensive epidemiological review. If empirical evidence is lacking, elicitation of experts' opinion would be a practical alternative.

4.5.4. Relative importance of risk factors and environmental descriptors

Decisions on the relative importance of the selected risk factors and environmental predictors which are used as proxies or surrogates for unmeasurable risk factors were informed by reviewing relevant literature as well as by obtaining experts' opinions.

We compared the results from AHP with those from other weight elicitation methods. AHP approach agreed with ranking order centroid method and swing weights technique, though results were different from balance beam approach (Edwards & Barron, 1994; Jia, Fischer, & Dyer, 1998; Watson & Buede, 1987) (Figure 4-19).
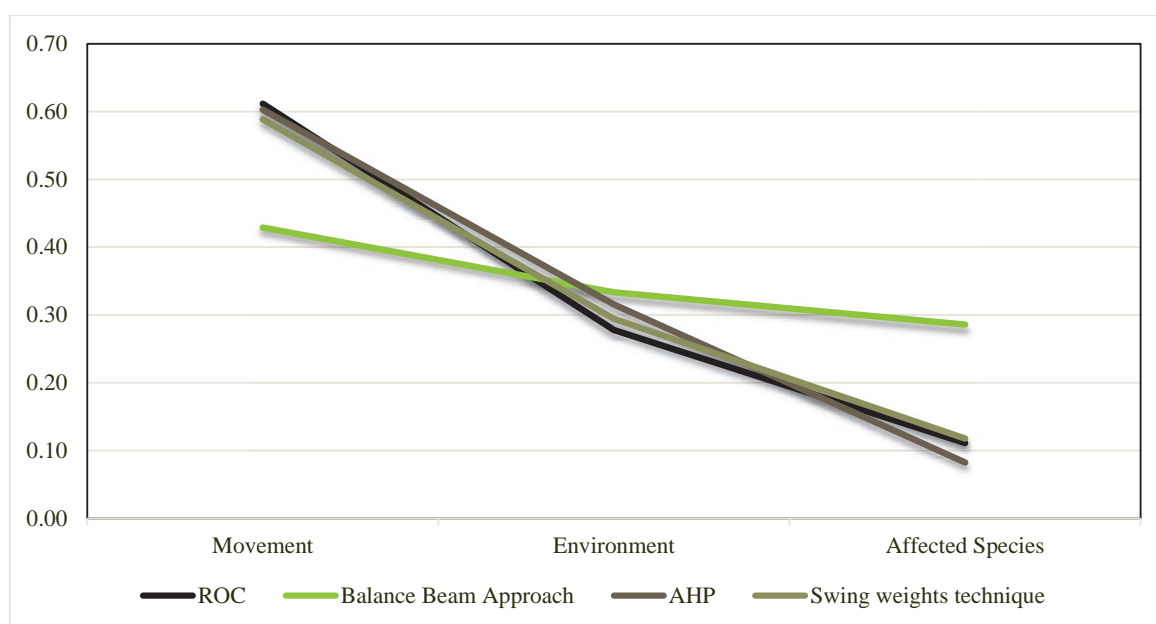
Figure 4-19. Comparison of results from weight elicitation methods for M-map
Note: Y-axis is weight value, with a range from 0 to 1. The larger the weight, the higher the value; ROC stands for Ranking Order Centroid and AHP stands for Analytic Hierarchy Process; Movement, Environment and Affected Species are three risk groups for HPAI H5N1.

4.5.5. Challenges and limitations

A risk map can be the outcome of modelling disease transmission, based on spatial and temporal data (Kitron, 2000), or can be directly derived from synthesis of available epidemiological information, as in HandiMap. A risk map attempts to represent ecological, biological, behavioral processes influenced by host, pathogen and surrounding environment and climate (Slingenbergh, 2004; Karesh, Dobson et al., 2012; Wilcox & Gubler, 2005). Such a process may misrepresent the dynamic process for any given disease if, for example, (i) the input data are outdated, or simply wrong; (ii) the explanatory variables being used are selected wrongly; (iii) the epidemiological assumption substantiating risk map is wrong (Hirzel, Hausser et al., 2002; Rogers & Randolph, 2003; Woolhouse, 2011). Besides, a pathogen or a disease will not occupy all suitable habitats. Hence even with the power of combining geographical information systems, remote sensing, computer technology and statistical tools, risk maps developed by using disease risk factors and their surrogates, individually or collectively can only estimate the likelihood of disease introduction or spread at a particular geographic location, and in some cases can also take account of time and season.

The risk maps produced by HandiMap are somewhat different in that they are intended as the starting point of an information gathering process, rather than the endpoint. To improve the predictability of any risk map, it is recommended that validation with disease presence data, either through ground truthing or

historical records and constant updating are undertaken (Brownstein, Skelly et al., 2005; Glass, Cheek et al., 2000; Sumption, Rweyemamu et al., 2008; Yang, Vounatsou et al., 2005). What HandiMap does is to integrate the best available existing information on the epidemiology of a disease and its likely spatial distribution, then use that through the modelling process to design a surveillance strategy to gather targeted information to enhance the accuracy of understanding of a disease and its possible presence in a country. This can be a cyclic process, as illustrated in Chapter 5, where development of a risk map led to risk-based surveillance which demonstrated that a disease was unexpectedly present in Mongolia, which then raised a list of questions that need to be answered to clarify the significance of the initial findings and to decide what action is needed.

Availability of proper data is a constant challenge in epidemiological investigation of emerging diseases. However, more and more public domain data on environmental and climatic descriptors relevant to disease risk mapping are now being generated from remote sensing images, which are global and can be readily manipulated and accessed. Food and Agriculture Organization of United Nations has also developed georeferenced data on numbers of major domestic animal species and economic crops. However these data are not always up to date or might not be available at the required resolution. Besides, some critical pieces of information such as level of disease control efforts by geographic unit or epidemiological unit, one of the important factors for disease risk estimation, have to be collected from the field.

Data availability is even more challenging in resource poor countries because of suboptimal or non-existing health information systems and public health programs. In such cases, creative ways have to be thought about how to use informative surrogates to approximate missing or unavailable data. For instance, one example was to use satellite nighttime light density images to estimate human population density (Pozzi, Small, & Yetman, 2003; Sutton, 2003). Other examples include using normalized difference vegetation index (NDVI) to estimate distribution of animal species (Leyequien et al., 2007) and tick abundance (Estrada-Peña, 1999).

HandiMap has been designed as a convenient disease risk mapping tool for infectious diseases. It is not intended to be a fully-fledged GIS software program to compete with commercial products such as Arc-GIS, or open source free products such as QGIS. It is specifically designed to be an integral part of the surveillance design function within HandiResponse, and it has now been shown to be capable of achieving that objective. The tool can be further improved in its user-friendliness.

4.6. Conclusions

HandiMap is a convenient tool specialized for estimating and visualizing the risk of occurrence for a given infectious disease. In the HandiMap module, infectious diseases are categorized into epitypes and mapping templates are available as examples for some of the epitypes, with the remaining examples to be developed after the end of this project. A range of environmental descriptor layers for constructing risk maps are stored on a server and can be used for disease risk mapping. This in turn leads to the design of epidemiologically and economically optimal disease surveillance programs, and potentially to more effective disease control.

It can facilitate effective communication on infectious diseases and decision making for surveillance and disease interventions, and it offers particular benefits in resource and data-sparse environments.

4.7. References

Balabanova, Y., A. Gilsdorf, et al. (2011). "Communicable diseases prioritized for surveillance and epidemiological research: results of a standardized prioritization procedure in Germany, 2011." Plos One 6(10): e25691.

Brownstein, J. S., C. C. Freifeld, et al. (2009). "Digital disease detection—harnessing the Web for public health surveillance." New England Journal of Medicine 360(21): 2153-2157.

Brownstein, J. S., D. K. Skelly, et al. (2005). "Forest fragmentation predicts local scale heterogeneity of Lyme disease risk." Oecologia 146(3): 469-475.

Cao, C., M. Xu, et al. (2010). "Risk analysis for the highly pathogenic avian influenza in Mainland China using meta-modelling." Chinese Science Bulletin 55(36): 4168-4178.

Cardoen, S., X. Van Huffel, et al. (2009). "Evidence-based semiquantitative methodology for prioritization of foodborne zoonoses." Foodborne Pathogens and Disease 6(9): 1083-1096.

Desvaux, S., V. Grosbois, et al. (2011). "Risk factors of highly pathogenic avian influenza H5N1 occurrence at the village and farm levels in the Red River Delta Region in Vietnam." Transboundary and emerging diseases 58(6): 492-502.

Doherty, J.-A. (2000). "Establishing priorities for national communicable disease surveillance." The Canadian Journal of Infectious Diseases 11(1): 21.

Drobne, S. and A. Lisec (2009). "Multi-attribute decision analysis in GIS: weighted linear combination and ordered weighted averaging." Informatica 33(4).

Eastman, J. R. (1997). Idrisi Para Windows: Version 2.0, Enero 1997. Tutorial Exercises, Clark University.

Edwards, W. and F. H. Barron (1994). "SMARTS and SMARTER: Improved simple methods for multiattribute utility measurement." Organizational behavior and human decision processes 60(3): 306-325.

Estrada-Peña, A. n. (1999). "Geostatistics and remote sensing using NOAA-AVHRR satellite imagery as predictive tools in tick distribution and habitat suitability estimations for Boophilus microplus (Acari: Ixodidae) in South America." Veterinary Parasitology 81(1): 73-82.

Fang, L.-Q., S. J. de Vlas, et al. (2008). "Environmental factors contributing to the spread of H5N1 avian influenza in mainland China." PloS one 3(5): e2268.

Gauthier - Clerc, M., C. Lebarbenchon, et al. (2007). "Recent expansion of highly pathogenic avian influenza H5N1: a critical review." Ibis 149(2): 202-214.

Ge, E., R. Haining, et al. (2012). "Using knowledge fusion to analyze avian influenza H5N1 in East and Southeast Asia."

Gilbert, M., P. Chaitaweesup, et al. (2006). "Free-grazing ducks and highly pathogenic avian influenza, Thailand." Emerging infectious diseases 12(2): 227-234.

Gilbert, M., S. H. Newman, et al. (2010). "Flying over an infected landscape: distribution of highly pathogenic avian influenza H5N1 risk in South Asia and satellite tracking of wild waterfowl." EcoHealth 7(4): 448-458.

Gilbert, M. and D. U. Pfeiffer (2012). "Risk factor modelling of the spatio-temporal patterns of highly pathogenic avian influenza (HPAI) H5N1: a review." Spatial and spatio-temporal epidemiology 3(3): 173-183.

Gilbert, M., X. Xiao, et al. (2007). "Avian influenza, domestic ducks and rice agriculture in Thailand." Agriculture, ecosystems & environment 119(3): 409-415.

Gilbert, M., X. Xiao, et al. (2008). "Mapping H5N1 highly pathogenic avian influenza risk in Southeast Asia." Proceedings of the National Academy of Sciences 105(12): 4769-4774.

Gilsdorf, A. and G. Krause (2011). "Prioritisation of infectious diseases in public health: feedback on the prioritisation methodology, 15 July 2008 to 15 January 2009."

Glass, G. E., J. E. Cheek, et al. (2000). "Using remotely sensed data to identify areas at risk for hantavirus pulmonary syndrome." Emerging infectious diseases 6(3): 238.

Havelaar, A. H., F. Van Rosse, et al. (2010). "Prioritizing emerging zoonoses in the Netherlands." Plos One 5(11): e13965.

Hay, S. I., K. E. Battle, et al. (2013). "Global mapping of infectious disease." Philosophical Transactions of the Royal Society of London B: Biological Sciences 368(1614): 20120250.

Henning, J. and D. U. Pfeiffer (2009). "Risk factors and characteristics of H5N1 Highly Pathogenic Avian Influenza (HPAI) post-vaccination outbreaks." Veterinary research 40(3): 1.

Henning, K. A., J. Henning, et al. (2009). "Farm-and flock-level risk factors associated with Highly Pathogenic Avian Influenza outbreaks on small holder duck and chicken farms in the Mekong Delta of Viet Nam." Preventive veterinary medicine 91(2): 179-188.

Hirzel, A. H., J. Hausser, et al. (2002). "Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data?" Ecology 83(7): 2027-2036.

Hogerwerf, L., R. G. Wallace, et al. (2010). "Persistence of highly pathogenic avian influenza H5N1 virus defined by agro-ecological niche." EcoHealth 7(2): 213-225.

Iglesias, I., M. J. Muñoz, et al. (2010). "Environmental risk factors associated with H5N1 HPAI in Ramsar wetlands of Europe." Avian diseases 54(2): 814-820.

Slingenbergh, M. G., K. de Balogh and W. Wint (2004). "Ecological sources of zoonotic diseases." Rev. sci. tech. Off. int. Epiz., 23(2): 467-484.

Jia, J., G. W. Fischer, et al. (1998). "Attribute weighting methods and decision quality in the presence of response error: a simulation study." Journal of Behavioral Decision Making 11(2): 85-105.

Karesh, W. B., A. Dobson, et al. (2012). "Ecology of zoonoses: natural and unnatural histories." The Lancet 380(9857): 1936-1945.

Kitron, U. (2000). "Risk maps: transmission and burden of vector-borne diseases." Parasitol Today 16(8): 324-325.

Krause, G. (2008). "How can infectious diseases be prioritized in public health?" EMBO reports 9(1S): S22-S27.

Kurowicka, D., C. Bucura, et al. (2010). "Probabilistic inversion in priority setting of emerging zoonoses." Risk Analysis 30(5): 715-723.

Lai, S. and L. Hopkins (1989). "The meanings of trade-offs in multiattribute evaluation methods: a comparison." Environment and Planning B 16(2): 155-170.

Leyequien, E., J. Verrelst, et al. (2007). "Capturing the fugitive: Applying remote sensing to terrestrial animal distribution and diversity." International Journal of Applied Earth Observation and Geoinformation 9(1): 1-20.

Loth, L., M. Gilbert, et al. (2011). "Identifying risk factors of highly pathogenic avian influenza (H5N1 subtype) in Indonesia." Preventive veterinary medicine 102(1): 50-58.

Malczewski, J. (1999). GIS and multicriteria decision analysis, John Wiley & Sons.

Mangen, M. J. J., M. B. Batz, et al. (2010). "Integrated approaches for the public health prioritization of foodborne and zoonotic pathogens." Risk Analysis 30(5): 782-797.

Martin, V., D. U. Pfeiffer, et al. (2011). "Spatial distribution and risk factors of highly pathogenic avian influenza (HPAI) H5N1 in China." PLoS Pathog 7(3): e1001308.

McKenzie, J., H. Simpson, et al. (2007). "Development of methodology to prioritise wildlife pathogens for surveillance." Preventive Veterinary Medicine 81(1): 194-210.

Ossadnik, W. and O. Lange (1999). "AHP-based evaluation of AHP-Software." European journal of operational research 118(3): 578-588.

Paul, M., S. Tavornpanich, et al. (2010). "Anthropogenic factors and the risk of highly pathogenic avian influenza H5N1: prospects from a spatial-based model." Veterinary research 41(3): 28.

Paul, M., S. Wongnarkpet, et al. (2011). "Risk factors for highly pathogenic avian influenza (HPAI) H5N1 infection in backyard chicken farms, Thailand." Acta tropica 118(3): 209-216.

Peterson, A. T. and R. A. Williams (2008). "Risk mapping of highly pathogenic avian influenza distribution and spread." Ecology and Society 13: 15.

Pfeiffer, D. U., P. Q. Minh, et al. (2007). "An analysis of the spatial and temporal patterns of highly pathogenic avian influenza occurrence in Vietnam using national surveillance data." The Veterinary Journal 174(2): 302-309.

Pigott, D. M., R. E. Howes, et al. (2015). "Prioritising Infectious Disease Mapping." PLoS Negl Trop Dis 9(6): e0003756.

Pires, S. M., E. G. Evers, et al. (2009). "Attributing the human disease burden of foodborne infections to specific sources." Foodborne Pathogens and Disease 6(4): 417-424.

Pozzi, F., C. Small, et al. (2003). "Modelling the distribution of human population with nighttime satellite imagery and gridded population of the world." Earth Observation Magazine 12(4): 24-30.

Rivas, A. L., G. Chowell, et al. (2010). "Lessons from Nigeria: the role of roads in the geo-temporal progression of avian influenza (H5N1) virus." Epidemiology and infection 138(02): 192-198.

Rogers, D. J. and S. E. Randolph (2003). "Studying the global distribution of infectious diseases using GIS and RS." Nature Reviews Microbiology 1(3): 231-237.

Si, Y., T. Wang, et al. (2010). "Environmental factors influencing the spread of the highly pathogenic avian influenza H5N1 virus in wild birds in Europe." Ecol. Soc 15: 26.

Sturm-Ramirez, K., D. Hulse-Post, et al. (2005). "Are ducks contributing to the endemicity of highly pathogenic H5N1 influenza virus in Asia?" Journal of virology 79(17): 11269-11279.

Sumption, K., M. Rweyemamu, et al. (2008). "Incidence and Distribution of Foot‐and‐Mouth Disease in Asia, Africa and South America; Combining Expert Opinion, Official Disease Information and Livestock Populations to Assist Risk Assessment." Transboundary and emerging diseases 55(1): 5-13.

Sutton, P. C. (2003). "14 Estimation of human population parameters using night-time satellite imagery." Remotely-Sensed Cities: 301.

Takekawa, J. Y., S. H. Newman, et al. (2010). "Migration of waterfowl in the East Asian flyway and spatial relationship to HPAI H5N1 outbreaks." Avian diseases 54(s1): 466-476.

Tatem, A. J., D. L. Smith, et al. (2010). "Ranking of elimination feasibility between malaria-endemic countries." The Lancet 376(9752): 1579-1591.

Tiensin, T., S. S. U. Ahmed, et al. (2009). "Ecologic risk factor investigation of clusters of avian influenza A (H5N1) virus infection in Thailand." Journal of Infectious Diseases 199(12): 1735-1743.

Tiensin, T., P. Chaitaweesup, et al. (2005). "Highly pathogenic avian influenza H5N1." Emerging infectious diseases 11: 1664-1672.

Van Boeckel, T. P., W. Thanapongtharm, et al. (2012). "Improving risk models for avian influenza: the role of intensive poultry farming and flooded land during the 2004 Thailand epidemic."

Van Boeckel, T. P., W. Thanapongtharm, et al. (2012). "Predicting the distribution of intensive poultry farming in Thailand." Agriculture, ecosystems & environment 149: 144-153.

Voogd, H. (1983). Multicriteria evaluation for urban and regional planning, Taylor & Francis.

Watson, S. R. and D. M. Buede (1987). Decision synthesis: the principles and practice of decision analysis, Cambridge University Press.

Wilcox, B. A. and D. J. Gubler (2005). "Disease ecology and the global emergence of zoonotic pathogens." Environmental Health and Preventive Medicine 10(5): 263-272.

Williams, R. A. and A. T. Peterson (2009). "Ecology and geography of avian influenza (HPAI H5N1) transmission in the Middle East and northeastern Africa." International Journal of Health Geographics 8(1): 47.

Woolhouse, M. (2011). "How to make predictions about future infectious disease risks." Philosophical Transactions of the Royal Society B: Biological Sciences 366(1573): 2045-2054.

Xiao, X., S. Boles, et al. (2006). "Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images." Remote Sensing of Environment 100(1): 95-113.

Xiao, X., S. Boles, et al. (2005). "Mapping paddy rice agriculture in southern China using multi-temporal MODIS images." Remote Sensing of Environment 95(4): 480-492.

Xiao, X., M. Gilbert, et al. (2007). "Remote sensing, ecological variables, and wild bird migration related to outbreaks of highly pathogenic H5N1 avian influenza." Journal of wildlife diseases 43(1): 40.

Yang, G.-J., P. Vounatsou, et al. (2005). "A review of geographic information system and remote sensing with applications to the epidemiology and control of schistosomiasis in China." Acta tropica 96(2): 117-129.

Yupiana, Y., S. J. de Vlas, et al. (2010). "Risk factors of poultry outbreaks and human cases of H5N1 avian influenza virus infection in West Java Province, Indonesia." International Journal of Infectious Diseases 14(9): e800-e805.

131

132

5. Evaluating the benefit of risk mapping to predict temporal and spatial distribution of Highly Pathogenic Avian Influenza (H5N1) outbreaks in southern Vietnam

5.1. Abstract

Background*:* Identification of a dynamic model which can accurately represent the spatial and temporal pattern of a disease outbreak is critical for development of an appropriately targeted surveillance program for a zoonotic disease.

Methods: A spatial and temporal model delineating the HPAI H5N1 outbreak in Southern Vietnam between late 2004 and early 2005 was developed by combining stochastic modelling, risk mapping and statistical tests. The stepwise approach comprises three key elements: (i) development of a model which provided statistically accurate temporal fit; (ii) using risk mapping to improve spatial fit of the base models developed in step 1; (iii) testing the statistical validity of the predictions by applying the Kolmogorov-Smirnov test for temporal fit, and Receiver-Operator Characteristic curves for spatial fit.

Results: The model which did not use risk mapping could predict temporal spread pattern adequately, but had only moderate capability to predict spatial spread. Risk mapping to provide a spatially variable landscape on which modelling of disease spread would take place could improve spatial prediction by this model, particularly when using a combination of using different disease risk factors and environmental predictors.

Conclusion: The approach was demonstrated to be a practical and generalizable method, which has the potential to be applied in other countries and for other diseases.

5.2. Introduction

Since HPAI H5N1 was first identified in East Asia in the mid-1990s (Duan, Bahl et al., 2008; Sims & Brown, 2008), it has quickly evolved into a pandemic disease affecting both animals and human beings in Asia, Europe and Africa (Adegboye & Kotze, 2014; Bett, Henning et al., 2014; Gauthier‑Clerc, Lebarbenchon, & Thomas, 2007; Hagag, Mansour et al., 2015; Stevens, Costard et al., 2009; Williams & Peterson, 2009). The disease has resulted in significant economic loss to rural households, to the wider agriculture sector and to the macro economy of affected countries. No previous true avian influenza virus strain had caused serious human disease, so the direct effect plus the negative externalities for the human population in the affected countries caused serious economic and social disruption, and rapid spread of the disease to many countries caused global panic and fear (Herring & Lockerbie, 2010; Osterholm, 2005; Peiris, De Jong, & Guan, 2007). The number of human cases has remained fairly small, and no sustained human-to-human transmission has yet occurred. Although it is not possible to predict if or when this may happen, it would only take a small number of critical genetic mutations or genomic replacements before the virus could be efficiently transmitted through human to human contact and flare up into a global human influenza pandemic (Amendola, Ranghiero et al., 2015; Belshe, 2005; Gambaryan & Matrosovich, 2015).

Poultry production in Vietnam is strategically important. Its productive value ranks second in the animal husbandry industry in Vietnam, and it has been steadily increasing since 1990s (Duc & Long, 2008) with the exception of the period between 2003 and 2005. Almost 80% of rural households participate in poultry production through backyard and garden rearing of ducks and/or chickens. Four regions which host the most poultry production in Vietnam are Red River Delta, South East Region, Mekong River Delta and South Central Coast Region (Burgos, Hanh et al., 2007).

Vietnam was among the hardest hit countries by HPAI H5N1. For instance, fifty nine out of sixty four provinces reported HPAI H5N1 outbreaks in 2007 (Minh, Morris et al., 2009). The human H5N1 cases reported from Vietnam accounted for more than 80 percent of the total cases around the world by 2007 (Beigel et al., 2005; Uyeki, 2008). The disease in the country has occurred as a series of epidemic waves since 2003, with gradually reducing size over the years as control measures took effect (Minh, Morris et al., 2009; Uyeki, 2008). From 2003 to 2007, the peak time of the reported human infections and poultry disease outbreaks shifted gradually from the initial time period of the cool season of December until March or April, to the warmer months of each year (Minh, Morris et al., 2009). The principal disease foci were in two geographically distinct ecological niches - around the Mekong River Delta in the south and the Red River Delta in the north. In these two epidemic zones the disease has become endemic, but there

are different epidemiological underpinnings in the two regions (Desvaux, Grosbois et al., 2011; Henning & Henning et al., 2009; Minh, Morris et al., 2009; Minh, Stevenson et al., 2010).

Models have become valuable tools and been increasingly integrated into the public health decision-making process. They can help to provide general insights on disease spread, generate testable hypotheses, predict disease outbreaks and assess the effect of control measures such as surveillance and vaccination campaigns (Barnabas, Laukkanen et al., 2006; Brisson, Edmunds et al., 2000; Keeling, Woolhouse et al., 2001; Koopman, 2005; Merler, Ajelli et al., 2015; Mossong, Hens et al., 2008; Myers, Rogers et al., 2000; Riley, 2007; Woolhouse, 2011). Spatial and temporal modelling tends to be more practical and desirable since it has the potential to predict where and when a disease would spread. Such pieces of information are particularly important at early stages of emerging infectious disease epidemics when little knowledge has accumulated about disease behaviour. However, there remains a need to further develop and test models of emerging diseases and to build them into decision support systems (Bettencourt & Ribeiro, 2008; Kao, 2002; Lawson & Leimich, 2000).

The HandiResponse is a disease surveillance management program which has been designed to assist countries to respond to the threat of introduction of emerging diseases. It is embedded in the Integrated Real-time Information System (IRIS) for Animal and Human Health, a data management platform for managing human and animal infectious diseases developed as a precursor project to development of HandiResponse, and has special value for managing zoonoses. IRIS is given different names in its use for different countries and purposes, and in this context it is called HandiView. HandiResponse comprises four modules at the moment: (i) risk map development – HandiMap; (ii) surveillance portfolio development – HandiSurv; (iii) economic impact assessment – HandiEcon and (iv) surveillance optimization – OptiSurv. The program can facilitate risk mapping, assessment of surveillance approaches and development of an economically optimal surveillance portfolio.

In this chapter, we plan to test the value of the software through evaluating how well the simulated spatial and temporal disease models refined by the risk information generated by using HandiMap can represent the spatial and temporal spread of a disease in a particular environment, during an outbreak in a country where emerging diseases are likely to occur. We use the dataset on H5N1 avian influenza in southern Vietnam during the second epidemic wave in 2004/5 to test the power of predictability of those models. The dataset from the Southern Viet Nam is relatively free from data biases, and with detailed spatial and temporal information on both HPAI H5N1 positive and negative communes. A range of raster maps of the spatial distribution of various factors which have been considered to influence local risk of disease outbreaks is also available, to test the potential value of alternative risk maps as a basis for evaluating the

135

effect of modelling the disease on a risk landscape which varies the transmission probabilities according to the local risk level, rather than on a uniform "flat" landscape, as is commonly used for disease modelling.

5.3. Methods

5.3.1. Overall technical approach

Using HandiResponse to develop best fit model for HPAI H5N1 outbreak in southern Vietnam took the following steps: (i) risk map development; (ii) creation of a base model to represent the HPAI H5N1 outbreak in Southern Vietnam in late 2004 and early 2005 without considering spatial variation in risk; (iii) reconfiguring the base model by using the information on the risk level of each commune generated by risk mapping using three alternative risk datasets based on different epidemiological assumptions (Figure 5-1), and to compare the spatial and temporal fit of the different model predictions to the recorded individual affected communes over the period of the actual outbreak.
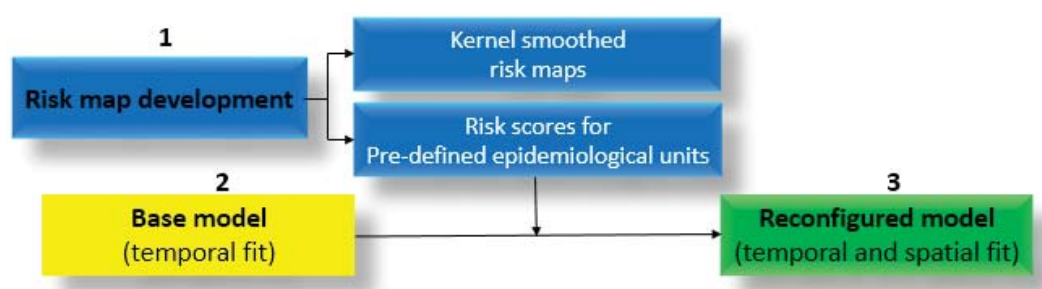
Figure 5-1. Approach for Screening best fit model for HPAI H5N1 outbreak in southern Vietnam

5.3.2. Risk map development

Three types of risk map for HPAI H5N1 spread in South Vietnam were developed to test alternative approaches: (i) environment-centered map, E-map; (ii) population-centred map, P-map and (iii) movement-centred map, M-map. Detailed procedure for HPAI H5N1 risk map development for southern Vietnam were documented in Chapter 4.

For each map developed, the value of each raster unit was based on the summation of weighted risk estimates for the particular one of the three risk assessment methods used. The raster surface was then kernel-smoothed, and the kernel-smoothed risk score attached to the centroid of each commune (the chosen epidemiological unit) was used to represent that commune, then the population of communes in southern Vietnam were divided into quintiles representing five risk levels. One of the five levels was

selected as the "reference level" and the same transmission probabilities were used for this level as in the base model. Then the transmission probabilities were scaled up or down for each of the other levels by adjustment factors which provided either a linear change in transmission probabilities or an escalating one, and different scaling options for the change in transmission probability were explored. Details of the adjustment process are shown in Figure 5-2. A total of 81 risk files (27 risk files for each of the three risk mapping strategies used) have been produced to allow comparison of the temporal and spatial fit of each risk-adjusted model with the base model.
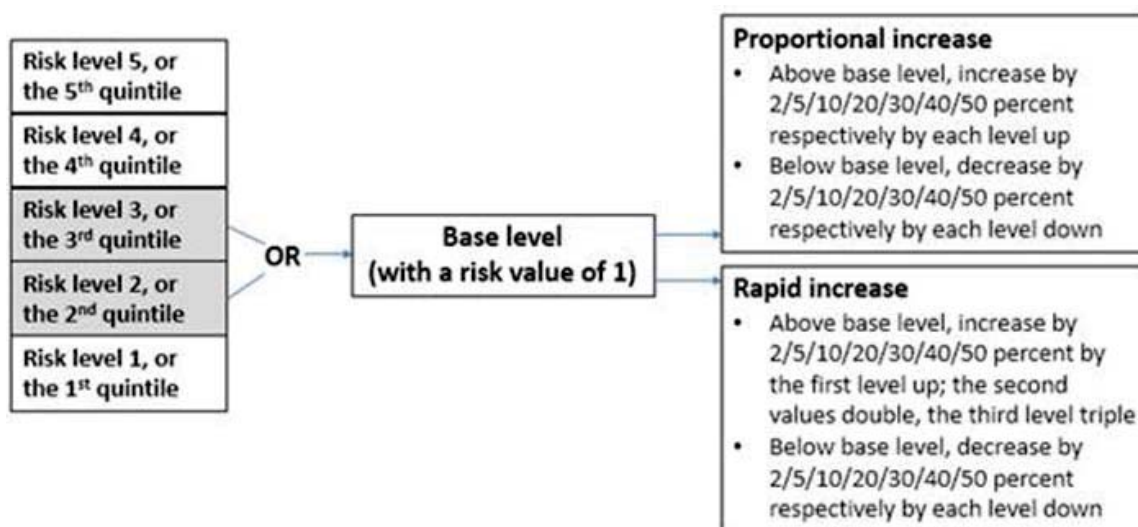
Figure 5-2. Adjustment procedure for transmission probabilities based on risk level

5.3.3.Development of avian influenza H5N1 outbreak model for Southern Vietnam

HandiSpread[4] was used to generate a model to represent the temporal and spatial characteristics of the avian influenza (HPAI) H5N1 epidemic wave in Southern Vietnam in late 2004 and early 2005. The HandiSpread software is a further development of the InterSpread Plus spatial disease model (Stevenson et al., 2013) to provide for modelling on risk landscape surfaces, and to include transfer of infection to human contacts of the infected animals and involvement of wildlife. Four different base models were developed and evaluated, with differences in important parameters to test which could give the best fit when modelling was conducted on a flat risk landscape with no risk factors taken into account. Then three alternative risk landscapes were derived from an analysis of the published literature – one based principally on environmental factors (E-Map), one based on population-related factors (P-Map), and one

---

[4] HandiSpread, Massey University EpiCentre, Palmerston North, New Zealand. Available at http://www.interspreadplus.com, accessed on August 14, 2014.

based on movement-related factors (M-Map). The corresponding spatial disease models shared the same initial parameter settings, but parameter values were adjusted at commune level to take account of the various risk maps, as described below, and the models were named respectively as E-model, P-model and M-model, corresponding to E-Map, P-Map and M-Map respectively, as shown in Figure 5-3. The procedures for developing the mapping process were described in Chapter 4.
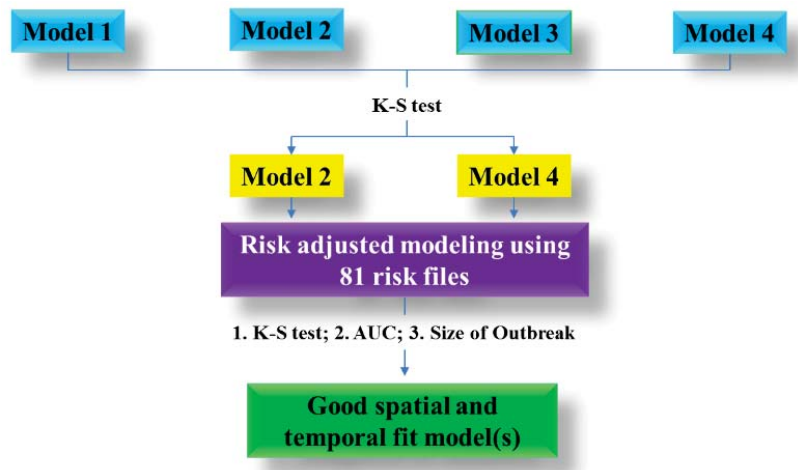
Figure 5-3. Plan of risk-adjusted disease modelling process.

Estimation of key parameters for base model development is summarized in Table 5-1.

Table 5-1. Key parameters needed for HPAI H5N1 outbreak model

Parameters	Values	Reference
Zone	South Vietnam (4 regions)	MARD, Vietnam
Epidemiological history	Day 1 (2 communes), day 2, 3 and 4 (1 each), day 5 (3)	Experts' estimation
Movement 1: live birds movement based on reported duck movement patterns	Number per period time: Poisson distribution, lambda is adjusted over the course of the epidemic wave - for model 1, 2 and 3 it takes values in the range from 0.001 to 0.68 but values differ, model 4 has values in the range 0.01 to 0.39 by time period. Distance and probability: 7 distance bands created between 0 km and 500 km. The probability of within band movement was assigned as following: Dis | 10 | 50 | 100 | 150 | 200 | 250 | 500 Pro | 0.2 | 0.4 | 0.08 | 0.05 | 0.05 | 0.18 | 0.04 Probability of Transmission: Constant 0.26	Experts' estimation Experts' estimation (Boender, Hagenaars et al., 2007)
Movement 2: poultry dealers	Number per period time: Poisson (Lambda=0.3) Distance and probability: 7 distance bands created between 0 km and 150 km. The probability of within band movement was assigned as following: Dis | 10 | 20 | 30 | 40 | 50 | 80 | 150 Pro | 0.65 | 0.1 | 0.1 | 0.08 | 0.03 | 0.03 | 0.01 Probability of Transmission: Constant 0.015	Experts' estimation Experts' estimation Experts' estimation
Movement 3: markets to farms	Number per time period: Constant 0 Distance and probability: 6 distance bands created between 0 km and 150 km. The probability of within band movement was assigned as following: Dis | 10 | 20 | 30 | 50 | 100 | 150 Pro | 0.25 | 0.22 | 0.2 | 0.15 | 0.1 | 0.08 Probability of Transmission: Constant 0.26	Experts' estimation Experts' estimation (Boender, Hagenaars et al., 2007)
Movement 4: farms to markets	Number per period time: Poisson distribution, Lambda = 0.14 Distance and probability: 6 distance bands created between 0 km and 150 km. The probability of within band movement was assigned as follows: Dis | 10 | 20 | 30 | 50 | 100 | 150 Pro | 0.5 | 0.25 | 0.1 | 0.05 | 0.09 | 0.01 Probability of Transmission: Constant 0.26	Expert's estimation Experts' estimation
Local spread	Probability of Transmission: 1 km (0.018), 2 km (0.01), 3 km (0.006) and 5 km (0.001)	(Boender, Hagenaars et al., 2007) (Mannelli, Ferre, & Marangon, 2006)
Infectivity	Time to clinical signs: Beta pert (6,8,13) Infectivity: Day | 0 | 1 | 2 | 3 | 6 | 13 | 16 | 17 Inf | 0 | 0.6 | 0.8 | 1 | 1 | 0.9 | 0.1 | 0	(Tiensin, Nielen et al., 2007) (Bos, Van Boven et al., 2007; Stegeman, Bouma et al., 2004; Van der Goot, Koch et al., 2005)

Note: Dis denotes distribution, Pro, probability, Inf, infectivity. Experts include Roger Morris and Bryan O'Leary.

5.3.4. Statistical analysis

SAS® (Version 9.4) was used for examining temporal and spatial fit of each model when compared with the actual outbreak in 2004 in Southern Vietnam. Kolmogorov-Smirnov two sample test (K-S Test) was performed to compare the temporal fit for epidemiological curve of daily cumulative number of infected communes between modelled outbreak and the actual outbreak. Area under the curve (AUC) of the receiver-operating characteristic (ROC) plot was used for assessing the spatial agreement between each model and the actual outbreak data at district level. Districts typically comprise 15-20 communes, with considerable variation. Although modelling used commune as the spatial unit for modelling transmission since that was the unit size at which outbreak data was collated. However it was too small a unit to be used to assess spatial fit of the model output to the reported epidemic. Further issues were the fact that there is extensive bird movement between communes, especially of grazing ducks, and hence commune is too small a unit for spatial fit comparisons, and also it is known that not all communes reported their disease outbreaks, whereas it is highly unlikely that an infected district would have no communes report signs consistent with H5N1 virus. Therefore district is considered a suitable unit size for spatial comparisons.

5.3.5. Identifying good fit model(s)

The following criteria were used for assessing model fit to the actual epidemic wave: temporal fit, better spatial fit than the base models and size of model outbreak (Table 5-2).

Table 5-2. Criteria for assessing fit of HPAI model(s) to actual epidemic data

Criterion	Explanation
Temporal fit	Epidemiological curve of daily cumulative infected communes is not significantly different from that of actual outbreak
Spatial fit	AUC of risk adjusted model should be larger than that of base model (calculated using AUC determined from the median plus 10 iterations on each side of the median iteration)
Size of outbreak	The modelled epidemic should provide an outbreak size at least as large as that of actual epidemic

5.4. Results

Of the four base models that vary in important parameters related to movement 1 evaluated, Models 2 and 4 gave the good temporal fit to the actual epidemic (Figure 5-4), and were chosen as the two to be used for tests and comparison after they were adjusted with the E-Map, P-Map and M-Map, as described before.

The three commune-level kernel-smoothed risk maps developed using the risk factors are presented in Figure 5-5, Figure 5-6 and Figure 5-7 below.
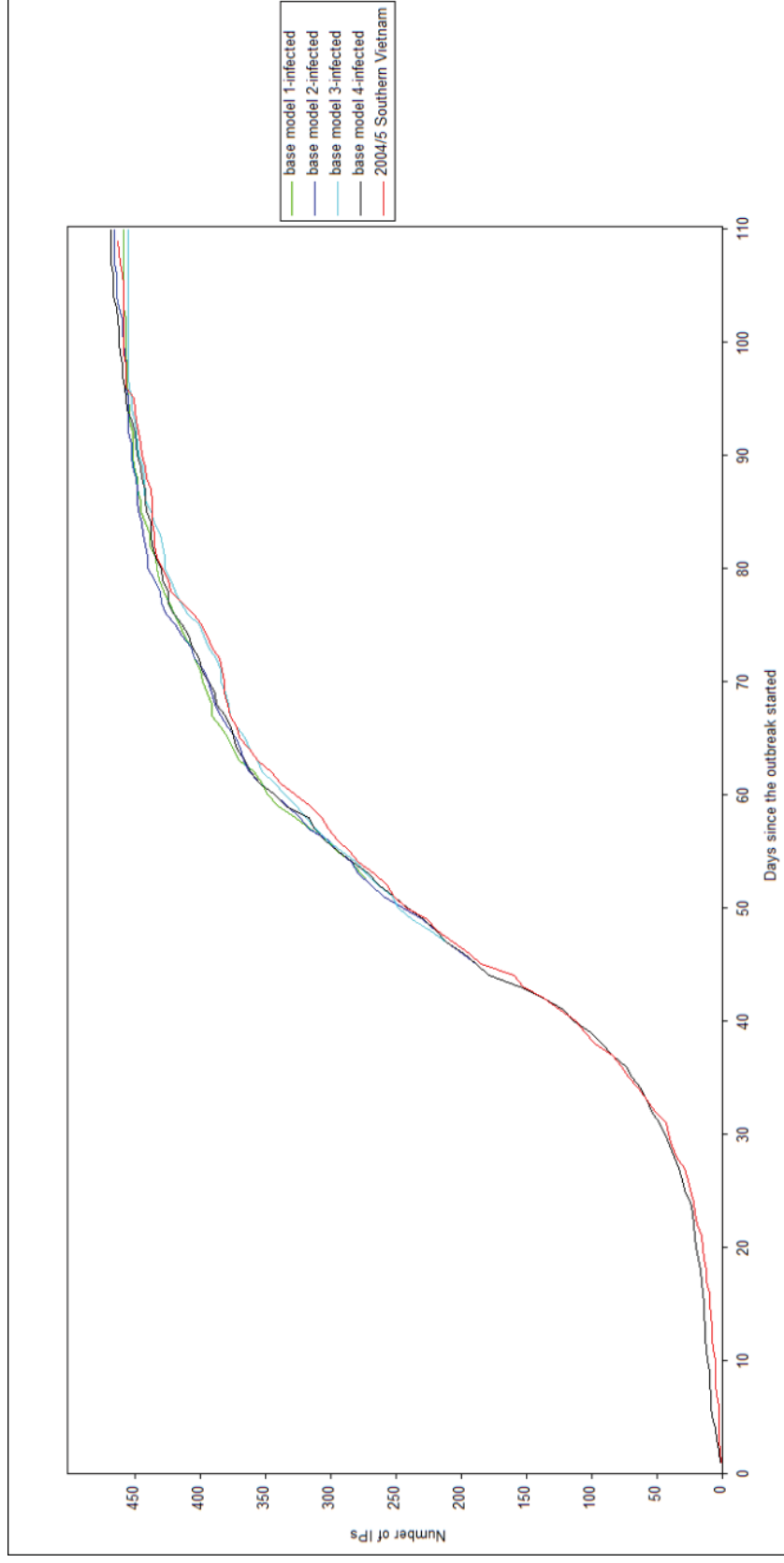
Figure 5-4. Comparison of four base models with the actual cumulative epidemic curve

Figure 5-5. Kernel smoothed E-map .

Figure 5-6. Kernel smoothed P-Map

Figure 5-7. Kernel smoothed M-map

Note: the darker the color, the higher the risk for HPAI H5N1 outbreak

143

Base models 2 and 4 were then simulated using the risk data generated from three risk maps derived from HandiResponse. Of the 162 alternative risk maps (81 maps each model) used, five were selected as the best ones meeting all three evaluation criteria, and these are summarized in Table 5-3. All these selected risk maps achieved close fit to the actual temporal epidemic curve, as demonstrated by the Kolmogorov-Smirnov test. All except one (m-r-3-5 in Table 5-3) had outbreak sizes larger than the reported outbreak. It is recognized that not all infected communes provided reports, so as long as the simulated total outbreak size for a model was equal to or moderately greater than the reported epidemic size, it would be considered to meet requirements. The main objective of the risk map was to improve spatial fit between model prediction and the actual epidemic, considering both positive and negative communes. The spatial fit was assessed by measuring the area under the ROC curve, reported as AUC. AUC was measured both for the single simulated outbreak of a median size (number of communes affected) (AUC2) and for the median plus ten simulated epidemics with the similar size to that of the median on either side of it (AUC1).

Four risk-adjusted models were found to be superior to Base Model 2 on spatial fit and equivalent on temporal fit. One model each used M-Map and P-Map, while two used E-Map. Model m-r-3-10 was considered the model which fitted best to the actual epidemic in all respects, when compared with base model 4. It is movement-centred, using risk level 3 as the reference level, with the risk level increasing by 10% to level 4 and 20% from level 4 to 5 (Table 5-3).

Table 5-3. Summary of goodness of fit of models to actual epidemic

Model	K-S Test	AUC 1	AUC 2	Size of outbreak
Base model 2	Accepted	0.63	0.60	466 (>462*)
p-r-3-5	Accepted	0.66	0.69	499
e-p-3-10	Accepted	0.65	0.68	468
e-r-3-5	Accepted	0.68	0.63	497
m-r-3-5	Accepted	0.65	0.72	459
Base model 4	Accepted	0.62	0.58	468
m-r-3-10	Accepted	0.69	0.74	480

Note: AUC 1, Average AUC for 21 iterations including median iteration and 10 iterations on each side of median iteration; AUC 2 denotes the value of AUC for median iteration. * the actual number of infected communes between late 2004 and early 2005 in southern Vietnam.

Spatial distributions of the actual outbreak, base model and the best risk adjusted model are presented in Figure 5-8. Temporal distributions of the actual outbreak and the good fit risk adjusted model are presented in Figure 5-9 and Figure 5-10.

6.a Actual outbreak 2004/5　　　　　6.b Base model 4　　　　　6.c Model 4-m-r-3-10
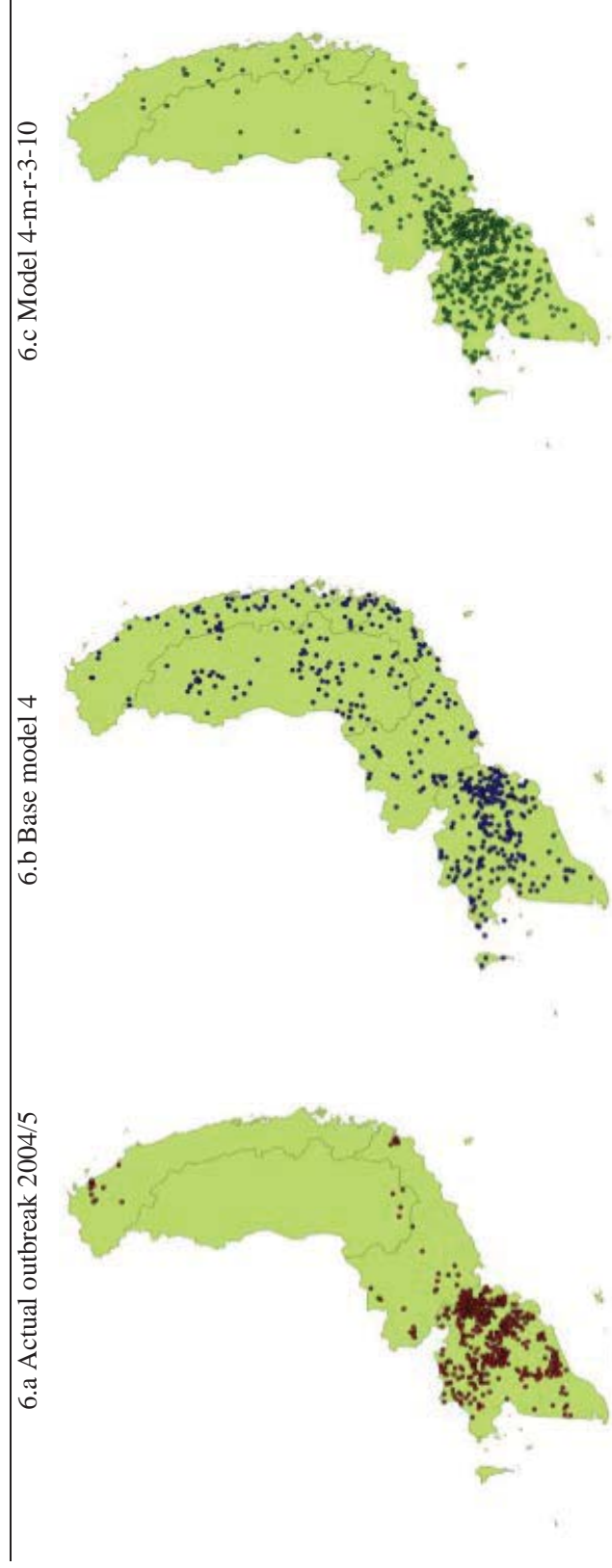
Figure 5-8. Comparison of spatial distribution of the good fit risk adjusted model with the actual outbreak in southern Vietnam, 2004/5

145
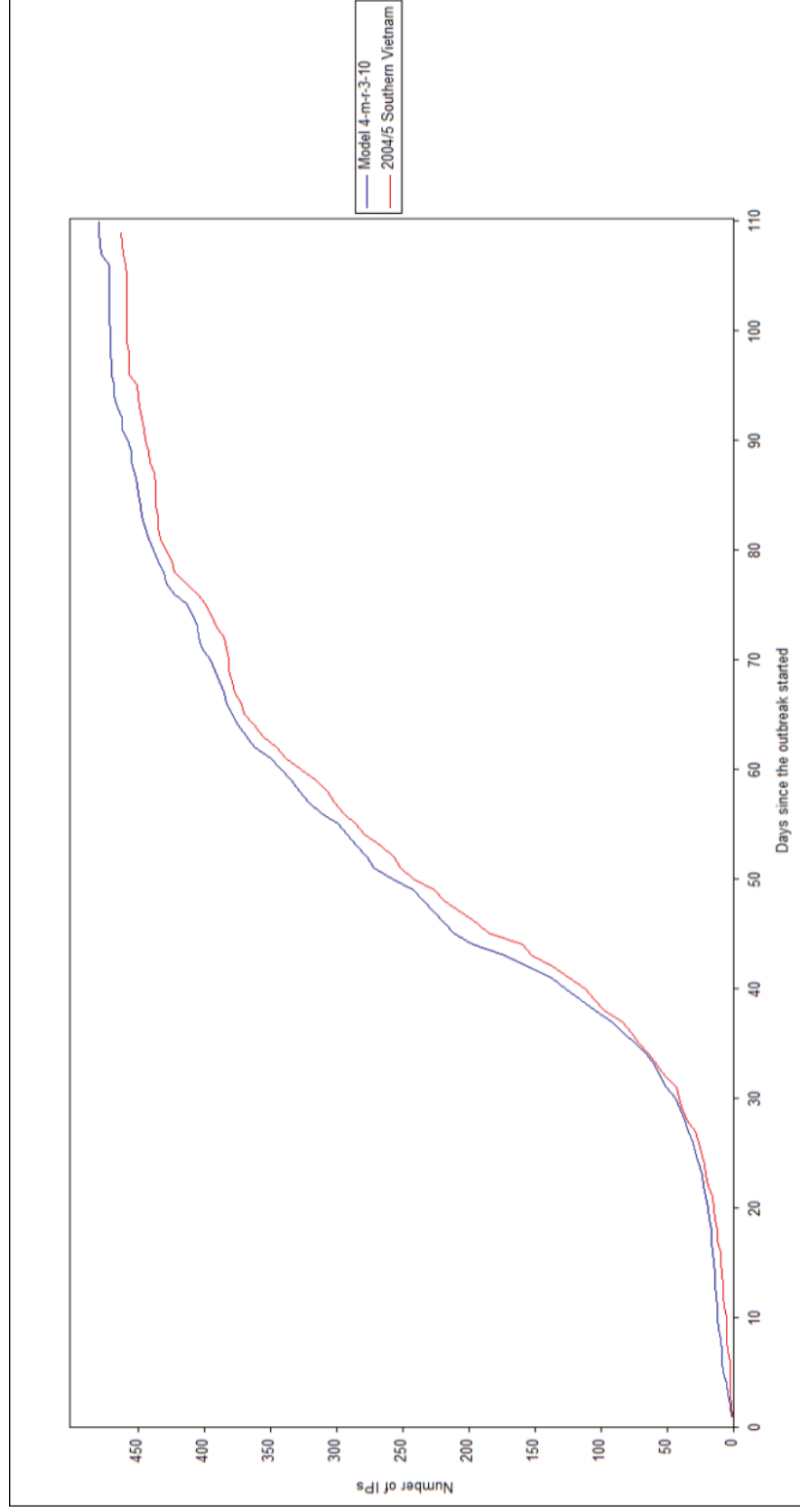
Figure 5-9. Comparison of daily cumulative number of infected communes of model 4-m-r-3-10 with that of the actual outbreak in Southern Vietnam, 2004/5
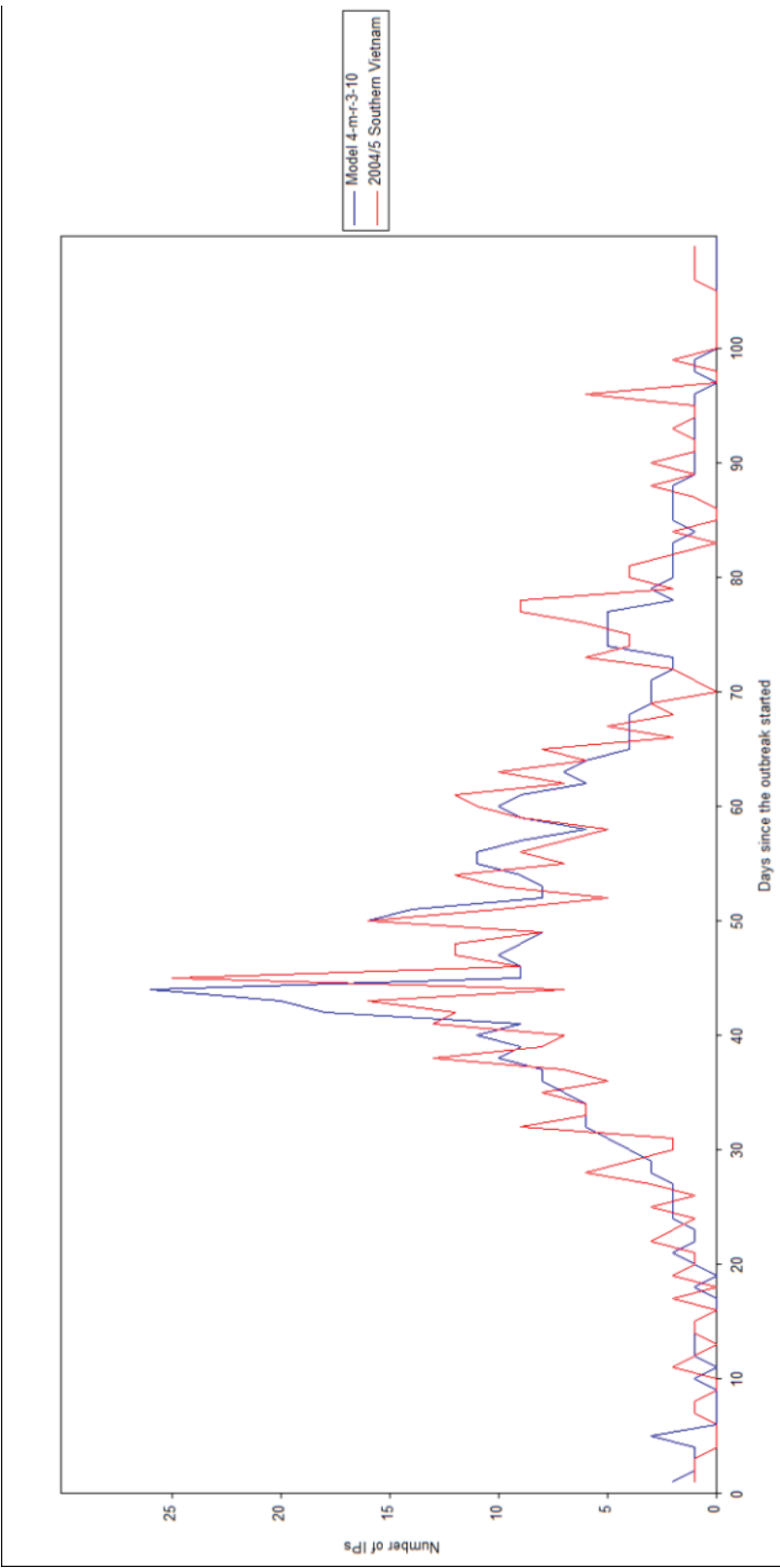
Figure 5-10. Comparison of number of daily infected communes under model 4-m-r-3-10 with that of the actual outbreak in Southern Vietnam, 2004/5

Kolmogorov-Smirnov test showed the temporal curve of model 4-m-r-3-10 was not significantly different from the actual HPAI outbreak in Southern Vietnam between late 2004 and early 2005 (Figure 5-11).

Kolmogorov-Smirnov Test for Variable Time Classified by Variable Group			
Group	N	EDF at Maximum	Deviation from Mean at Maximum
Base model	480	0.425000	0.845423
Actual 2004	462	0.346320	-0.861735
Total	942	0.386412	
Maximum Deviation Occurred at Observation 125			
Value of Time at Maximum = 44.0			

Kolmogorov-Smirnov Two-Sample Test (Asymptotic)			
KS	0.039333	D	0.078680
KSa	1.207198	Pr > KSa	0.1084

Figure 5-11. Result of Kolmogorov Smirnov Test

Results of AUC, generated from SAS Program, for the median iteration and 10 iterations on each side of the median iteration are documented in Table 5-4. The size of AUC (AUC 1) of the median iteration is 0.74 and the average size AUC (AUC2) of all 21 iterations is 0.69.

Table 5-4. ROC association statistics for Model 4-m-r-3-10

ROC	Area	Standard Error	Mann-Whitney 95% Wald Confidence Limits		Somers' D (Gini)	Gamma	Tau-a
1	0.7024	0.0242	0.655	0.7497	0.4047	0.7381	0.1867
2	0.7082	0.0253	0.6585	0.7578	0.4163	0.72	0.192
3	0.6627	0.0262	0.6113	0.7141	0.3254	0.6043	0.1501
4*	**0.7446**	**0.0237**	**0.6981**	**0.7911**	**0.4892**	**0.8094**	**0.2257**
5	0.6959	0.0253	0.6463	0.7456	0.3918	0.6971	0.1808
6	0.7119	0.025	0.6628	0.7609	0.4237	0.7336	0.1955
7	0.6847	0.0256	0.6345	0.7349	0.3694	0.6684	0.1704
8	0.6151	0.0251	0.5659	0.6643	0.2302	0.5012	0.1062
9	0.702	0.025	0.6531	0.7509	0.404	0.7175	0.1863
10	0.6963	0.0246	0.648	0.7446	0.3926	0.7153	0.1811
11	0.5971	0.0271	0.544	0.6502	0.1941	0.3921	0.0896
12	0.7048	0.0251	0.6556	0.754	0.4096	0.7194	0.189
13	0.6236	0.0258	0.5731	0.6742	0.2473	0.5089	0.1141
14	0.6362	0.0252	0.5869	0.6856	0.2725	0.5623	0.1257
15	0.7539	0.0225	0.7097	0.798	0.5078	0.8462	0.2342
16	0.7681	0.0233	0.7225	0.8137	0.5362	0.8436	0.2473
17	0.6256	0.0265	0.5737	0.6776	0.2512	0.4969	0.1159
18	0.7211	0.024	0.6741	0.7682	0.4423	0.7715	0.204
19	0.6902	0.0251	0.6411	0.7394	0.3805	0.693	0.1755
20	0.7456	0.024	0.6985	0.7927	0.4912	0.804	0.2266
21	0.7253	0.0238	0.6788	0.7719	0.4507	0.7836	0.2079
Average	**0.69**						

Note: The table was generated by SAS Program. Column One presents serial number for each individual iteration and the fourth iteration with * represents the median iteration.

149

5.5. Discussion

5.5.1. Rationale for the selection of Wave II HPAI H5N1 outbreak in Southern Vietnam

There are only a few data sets on disease epidemics in Asia where there is an adequate spatial data set of outbreak locations, and data on putative risk factors for the spread of the disease. One of these data sets exists for the avian influenza H5N1 epidemic in Vietnam, although even that has limitations that must be considered in interpretation of the evidence. The available data and publications derived from the data set were used to decide on the best sub-set of the total data set to be used in testing whether the use of a risk landscape could improve the predictive value of a spatial model. It was not realistic to develop a single set of model parameters to represent the HPAI H5N1 outbreaks in the whole of Vietnam. As informed by the relevant empirical studies, HPAI H5N1 outbreaks in Red River Delta and Mekong River Delta were similar in several respects but also epidemiologically divergent in a number of respects, driven to a significant extent by different risk factors. For example, whereas there was quite close temporal synchrony between the two regions in epidemic waves I and II, there was a substantial third epidemic wave in Red River Delta associated with the first vaccination campaign, but only a slight increase in outbreaks in Mekong Delta. It has been suggested that this was due to the riskier vaccination policy known to have been used in Red River Delta. Even where the same risk factors were operating, they might have different epidemiological importance (Minh, Morris et al., 2009; Pfeiffer, Minh et al., 2007). Development of a single risk map for Vietnam by assuming the same level of importance in south and north could result in misrepresentation, unless detailed epidemiological evidence could be used to adjust parameter settings.

The second wave of HPAI H5N1 outbreaks in the Mekong River delta region between December 2004 and April 2005 was used as the test case to evaluate how accurately the reported epidemiological curve for the disease outbreaks could be replicated in HandiSpread, using different risk landscapes. As depicted by Figure 5-12, the epidemic of Wave I HPAI H5N1 outbreak in 2003 was characterized as of shorter duration and with a peak around late January with an abrupt decrease afterwards that was unlikely to be real. The initial phase of this first epidemic wave is known to have been affected by under-reporting and problems in establishing adequate records. The rapid decrease in the late phase of the wave was partly explained by depopulation of many flocks and partly attributed to under-reporting (Gilbert, Xiao et al., 2008). On the other hand, the size of the late 2005 outbreak was much smaller and interpretation of the pattern is complicated by the progressive implementation of vaccination campaigns. The reported epidemic wave 2 in 2004/5 was considered to be closest to the behaviour of the true underlying epidemic because (i) disease reporting was better than in 2003 as farmers and veterinarians gained some knowledge

of the disease; (ii) complicating interventions such as vaccination campaigns were not introduced until later in 2005.
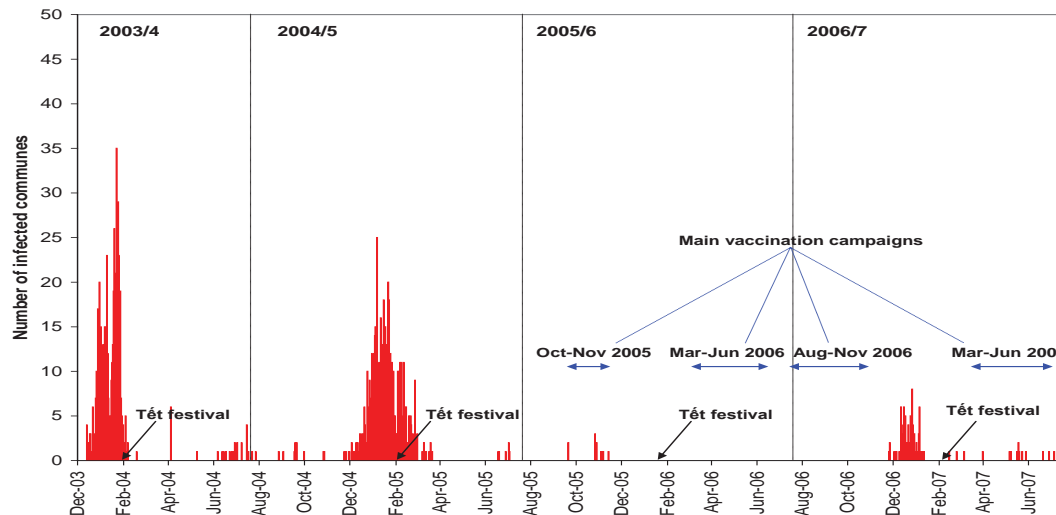
Figure 5-12. Epidemic Waves of HPAI H5N1 in Mekong River Delta, Vietnam
Note: with permission, the figure was adopted from the original article by Minh, P. Q., R. S. Morris, et al. (2009). "Spatio-temporal epidemiology of highly pathogenic avian influenza outbreaks in the two deltas of Vietnam during 2003–2007." Preventive veterinary Medicine 89(1): 16-24.

5.5.2. Identification of the best fit model(s)

Four potential candidates were initially developed as the base models. They differed from each other by the value for the parameter of "time (frequency) per period time" under Movement 1. Movement 1 was created to typify the itinerant duck movements that were hypothesized as the main risk factor in the South East Asia setting (Gaidet, Newman et al., 2008; Gilbert, Xiao et al., 2007; Kim, Negovetich et al., 2009). Eventually base model 2 and model 4 were chosen as the finalist models at this stage for further modelling works since they resembled the actual 2004/5 HPAI outbreak in Southern Vietnam temporally by having passed the Kolmogorov-Smirnov test.

In the second stage, the farm file used in base model 2 and 4 was replaced by each of the 81 risk adjusted farm files (27 risk files for each of the three risk mapping strategies) generated from three types of risk maps in an attempt to improve their spatial fit. A number of promising candidates, out of a total number of 162 variant models, were screened out for K-S test and calculation of AUC. Eventually the models with no significant temporal difference from the actual outbreak, with a similar or larger size of outbreak, and larger value of AUC were selected as the spatially best fitting candidates. The detailed technical procedure for selecting good fit model(s) is illustrated in previous Figure 5-3.

151

The study demonstrates that M-model, using a combination of different risk factors and environment predictors, represents the best option, followed by environmental predictors represented in E-model. The least satisfactory model was P-model.

5.5.3. Relative importance of risk factors and environmental predictors

We assumed that poultry movement played a slightly important role than environment predictors for the best available model, m4-m-r-3-10, based upon literature review and experts' opinion. We then tested three additional hypotheses: (i) movement was epidemiologically more important than environmental predictors in term of contribution to HPAI H5N1 outbreak; (ii) both of them were equally important and (iii) movement was less important than environment predictors (Table 5-5). However, compared with the m4-m-r-3-10, none of the three additional options improved the spatial fit (Table 5-6).

Table 5-5. Weight assignments under three additional hypotheses

Risk Factor	Hypothesis 1*	Hypothesis 2**	Hypothesis 3***
Distance to highway	0.60	0.39	0.13
Distance to waterway	0.12	0.08	0.07
Rice Paddy	0.18	0.41	0.64
% land covered by water	0.03	0.06	0.09
Duck Density	0.06	0.05	0.05
Chicken Density	0.01	0.01	0.01
Sum	**1.00**	**1.00**	**1.00**

Note: *movement is epidemiologically more important than environmental predictors; **movement is equally important as environmental predictors; ***movement is less important than environmental predictors.

Table 5-6. Average AUC under three additional hypotheses

Model	K-S Test	AUC 1	AUC 2	Size of outbreak
m4-2-m-r-3-4	Accepted	0.65	0.66	The size is larger than that for the actual outbreak (487>462)
m4-3-m-p-2-2	Accepted	0.64	0.63	The size is larger than that for the actual outbreak (470>462)
m4-4-m-p-2-2	Accepted	0.65	0.68	The size is less than that for actual outbreak (451<462)

Note: AUC 1, Average AUC for 21 iterations including median iteration and 10 iterations on each side of median iteration; AUC 2 denotes the value of AUC for median iteration.

5.5.4. The benefit of using a risk map for modelling H5N1 avian influenza

The objective in this study was to explore the effect of taking a map of risk factor information into account in predictive modelling of avian influenza H5N1 in Vietnam, and assessing whether doing this improves predictive accuracy of the model. The answer is that it improves the spatial prediction of the model, without sacrificing temporal fit (which was an inherent risk of the strategy).

This was a first attempt, and it will be necessary over time to broaden the range of disease and countries for which the approach is tested in order to more fully judge its potential, but that will require better data collection on disease events by countries, to provide a wider range of data sets. Complicating the assessment was the fact that there was no information on true positives and true negatives at commune level. The information used to represent the actual HPAI H5N1 outbreak was based on reporting that was highly likely to be compromised to a moderate degree because of imperfect knowledge of HPAI disease among farmers, veterinarians as well as the negative impacts such as banning of poultry trade, reduced consumption, culling of flocks with unfulfilled or suboptimal compensation provided, etc. If it had been possible to cross-check between clinical reports and laboratory investigation results on the same communes (both positive report communes and communes which did not report the disease), the predictive accuracy of the model(s) could possibly be further improved. Such an approach has been tested elsewhere (Martin, Pfeiffer, et al., 2011).

An important benefit of comparing model predictions with actual outbreaks is that it provides insights into the epidemiological processes on disease spread. Models generated by HandiSpread provide estimation of emergent properties. The emergent properties are the statistics of contributions to the outbreak by different transmission modes. Through comparing these emergent properties with field evidence, the predicted values of the models can be verified and models can be further improved. For instance, in this study, the modelling predicted that live bird markets and itinerant duck movement explained 49 percent and 40 percent of HPAI outbreaks respectively, and the contributions from other routes were insignificant (Table 5-7). Such pieces of information could be used for priority setting in intervention design.

Table 5-7. Epidemiological contribution of different transmission routes under m4-m-r-3-10

Mode of Infection	Number of IPs	Share
Epidemic History	8	2%
Local spread	34	7%
Live bird movement	194	40%
Personnel movement	9	2%
From live bird markets	235	49%
To live bird markets	0	0
Total	**480**	**100%**

Note: IP stands for infected property.

We discerned that daily reported number of outbreaks by communes was reduced from about 25 per day before Tet (New Year) Festival to less than 5 per day around the Tet Festival in early February, then increased sharply again after Tet. Although it could be partly explained by reduced poultry movement and

trade during this period of time, it would be hard to believe the reduction to be so dramatic. It is likely that this represented delayed reporting, and therefore the model should predict levels over Tet higher that the field data showed, which was in fact the case.

This study has revealed that the key to develop a good fit model is (i) to understand epidemiological processes driving the disease spread. For instance, it is critical to represent different movement types varying by distance, frequency and probability of transmission, plus local spread of disease in model development and (ii) to identify and represent key epidemiological units/groups and interaction between them.

Selection of the epidemiological unit for analysis was also critical to accurately judging the predictive power of models. While the unit of observation for disease outbreaks was the commune, this was too fine a spatial scale to use in assessing model performance, and the larger district unit was therefore used. This better represented the pattern of spread as predicted by the model, and reduced the effect of non-reporting by some communes.

5.5.5.Future direction

This exercise was the first attempt to use HandiResponse program to develop risk maps and create a dynamic spatial and temporal model for a disease which took account of spatial variation in susceptibility of local areas to disease establishment. It demonstrated that this approach improved the spatial fit of model predictions, without adversely affecting other aspects of prediction. As next steps, we intend to improve the user-friendliness of the program for users in countries at risk of disease outbreaks, and also to use the model to test different surveillance options.

5.6. Conclusion

Disease modelling can help provide general insights on disease spread, generate testable hypotheses, predict disease outbreak patterns and assess effect of control measures such as surveillance and vaccination campaigns. It can also facilitate decision making processes. This study demonstrated that a model which took account of risk maps could provide useful predictions of both temporal and spatial aspects of disease spread in a real disease outbreak, HPAI H5N1 epidemic wave II in southern Vietnam.

154

5.7. References

Adegboye, O. and D. Kotze (2014). "Epidemiological analysis of spatially misaligned data: a case of highly pathogenic avian influenza virus outbreak in Nigeria." Epidemiology and infection 142(05): 940-949.

Amendola, A., A. Ranghiero, et al. (2015). "Is avian influenza virus A (H5N1) a real threat to human health?" Journal of preventive medicine and hygiene 52(3).

Barnabas, R. V., P. Laukkanen, et al. (2006). "Epidemiology of HPV 16 and cervical cancer in Finland and the potential impact of vaccination: mathematical modelling analyses." PLoS Med 3(5): e138.

Beigel, J. H., J. Farrar, et al. (2005). "Avian influenza A (H5N1) infection in humans." New England Journal of Medicine 353(13): 1374-1385.

Belshe, R. B. (2005). "The origins of pandemic influenza—lessons from the 1918 virus." New England Journal of Medicine 353(21): 2209-2211.

Bett, B., J. Henning, et al. (2014). "Transmission rate and reproductive number of the H5N1 highly pathogenic avian influenza virus during the December 2005–July 2008 epidemic in Nigeria." Transboundary and emerging diseases 61(1): 60-68.

Bettencourt, L. M. and R. M. Ribeiro (2008). "Real time bayesian estimation of the epidemic potential of emerging infectious diseases." PLoS One 3(5): e2185.

Boender, G. J., T. J. Hagenaars, et al. (2007). "Risk maps for the spread of highly pathogenic avian influenza in poultry." PLoS Comput Biol 3(4): e71.

Bos, M. E., M. Van Boven, et al. (2007). "Estimating the day of highly pathogenic avian influenza (H7N7) virus introduction into a poultry flock based on mortality data." Veterinary research 38(3): 493-504.

Brisson, M., W. Edmunds, et al. (2000). "Modelling the impact of immunization on the epidemiology of varicella zoster virus." Epidemiology and infection 125(03): 651-669.

Burgos, S., P. H. Hanh, et al. (2007). "Characterization of poultry production systems in Vietnam." International Journal of Poultry Science 6(10): 709-712.

Desvaux, S., V. Grosbois, et al. (2011). "Risk factors of highly pathogenic avian influenza H5N1 occurrence at the village and farm levels in the Red River Delta Region in Vietnam." Transboundary and emerging diseases 58(6): 492-502.

Duan, L., J. Bahl, et al. (2008). "The development and genetic diversity of H5N1 influenza virus in China, 1996–2006." Virology 380(2): 243-254.

Duc, N. V. and T. Long (2008). "Poultry Production Systems in Viet Nam." Animal genetics and breeding department, National Institute of animal husbandry, Vietnam.

Gaidet, N., S. H. Newman, et al. (2008). "Duck migration and past influenza A (H5N1) outbreak areas." Emerging infectious diseases 14(7): 1164.

Gambaryan, A. and M. Matrosovich (2015). "What adaptive changes in hemagglutinin and neuraminidase are necessary for emergence of pandemic influenza virus from its avian precursor?" Biochemistry (Moscow) 80(7): 872-880.

Gauthier - Clerc, M., C. Lebarbenchon, et al. (2007). "Recent expansion of highly pathogenic avian influenza H5N1: a critical review." Ibis 149(2): 202-214.

Gilbert, M., X. Xiao, et al. (2007). "Avian influenza, domestic ducks and rice agriculture in Thailand." Agriculture, ecosystems & environment 119(3): 409-415.

Gilbert, M., X. Xiao, et al. (2008). "Mapping H5N1 highly pathogenic avian influenza risk in Southeast Asia." Proceedings of the National Academy of Sciences 105(12): 4769-4774.

Hagag, I. T., S. M. Mansour, et al. (2015). "Pathogenicity of Highly Pathogenic Avian Influenza Virus H5N1 in Naturally Infected Poultry in Egypt."

Henning, K. A., J. Henning, et al. (2009). "Farm-and flock-level risk factors associated with Highly Pathogenic Avian Influenza outbreaks on small holder duck and chicken farms in the Mekong Delta of Viet Nam." Preventive veterinary medicine 91(2): 179-188.

Herring, D. A. and S. Lockerbie (2010). "The Coming Plague of Avian Influenza." Plagues and Epidemics: Infected Spaces Past and Present: 179.

Kao, R. R. (2002). "The role of mathematical modelling in the control of the 2001 FMD epidemic in the UK." Trends in microbiology 10(6): 279-286.

Keeling, M. J., M. E. Woolhouse, et al. (2001). "Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape." Science 294(5543): 813-817.

Kim, J. K., N. J. Negovetich, et al. (2009). "Ducks: the "Trojan horses" of H5N1 influenza." Influenza and other respiratory viruses 3(4): 121-128.

Koopman, J. S. (2005). "Infection transmission science and models." Japanese journal of infectious diseases 58(6): S.

Lawson, A. B. and P. Leimich (2000). "Approaches to the space-time modelling of infectious disease behaviour." Mathematical Medicine and Biology 17(1): 1-13.

Mannelli, A., N. Ferre, et al. (2006). "Analysis of the 1999–2000 highly pathogenic avian influenza (H7N1) epidemic in the main poultry-production area in northern Italy." Preventive veterinary medicine 73(4): 273-285.

Martin, V., D. U. Pfeiffer, et al. (2011). "Spatial distribution and risk factors of highly pathogenic avian influenza (HPAI) H5N1 in China." PLoS Pathog 7(3): e1001308.

Merler, S., M. Ajelli, et al. (2015). "Spatiotemporal spread of the 2014 outbreak of Ebola virus disease in Liberia and the effectiveness of non-pharmaceutical interventions: a computational modelling analysis." The Lancet Infectious Diseases 15(2): 204-211.

Minh, P. Q., R. S. Morris, et al. (2009). "Spatio-temporal epidemiology of highly pathogenic avian influenza outbreaks in the two deltas of Vietnam during 2003–2007." Preventive veterinary medicine 89(1): 16-24.

Minh, P. Q., M. Stevenson, et al. (2010). "A description of the management of itinerant grazing ducks in the Mekong River Delta of Vietnam." Preventive veterinary medicine 94(1): 101-107.

Mossong, J., N. Hens, et al. (2008). "Social contacts and mixing patterns relevant to the spread of infectious diseases." PLoS Med 5(3): e74.

Myers, M. F., D. Rogers, et al. (2000). "Forecasting disease risk for increased epidemic preparedness in public health." Advances in Parasitology 47: 309-330.

Osterholm, M. T. (2005). "Preparing for the next pandemic." New England Journal of Medicine 352(18): 1839-1842.

Peiris, J. M., M. D. De Jong, et al. (2007). "Avian influenza virus (H5N1): a threat to human health." Clinical microbiology reviews 20(2): 243-267.

Pfeiffer, D. U., P. Q. Minh, et al. (2007). "An analysis of the spatial and temporal patterns of highly pathogenic avian influenza occurrence in Vietnam using national surveillance data." The Veterinary Journal 174(2): 302-309.

Riley, S. (2007). "Large-scale spatial-transmission models of infectious disease." Science 316(5829): 1298-1301.

Sims, L. D. and I. H. Brown (2008). "Multicontinental epidemic of H5N1 HPAI virus (1996–2007)." Avian influenza: 251-286.

Stegeman, A., A. Bouma, et al. (2004). "Avian influenza A virus (H7N7) epidemic in The Netherlands in 2003: course of the epidemic and effectiveness of control measures." Journal of Infectious Diseases 190(12): 2088-2095.

Stevens, K., S. Costard, et al. (2009). "Mapping the Likelihood of Introduction and Spread of Highly Pathogenic Avian Influenza Virus H5N1 in Africa, Ghana, Ethiopia, Kenya and Nigeria using Multicriteria Decision Modelling." DFID funded project for Controlling Avian Flu and Protecting People's Livelihoods in Africa/Indonesia.

Stevenson, M., R. Sanson, et al. (2013). "InterSpread Plus: a spatial and stochastic simulation model of disease in animal populations." Preventive Veterinary Medicine 109(1): 10-24.

Tiensin, T., M. Nielen, et al. (2007). "Transmission of the highly pathogenic avian influenza virus H5N1 within flocks during the 2004 epidemic in Thailand." Journal of Infectious Diseases 196(11): 1679-1684.

Uyeki, T. M. (2008). "Global epidemiology of human infections with highly pathogenic avian influenza A (H5N1) viruses." Respirology 13(s1): S2-S9.

Van der Goot, J., G. Koch, et al. (2005). "Quantification of the effect of vaccination on transmission of avian influenza (H7N7) in chickens." Proceedings of the National Academy of Sciences of the United States of America 102(50): 18141-18146.

Williams, R. A. and A. T. Peterson (2009). "Ecology and geography of avian influenza (HPAI H5N1) transmission in the Middle East and northeastern Africa." International Journal of Health Geographics 8(1): 47.

Woolhouse, M. (2011). "How to make predictions about future infectious disease risks." Philosophical Transactions of the Royal Society B: Biological Sciences 366(1573): 2045-2054.

6. HandiMap Assisted Risk-based Survey on Crimean Congo Haemorrhagic Fever in Mongolia

6.1. Abstract

Background: The ecological conditions and fauna in some parts of Mongolia are similar to those in surrounding countries, where Crimean Congo Haemorrhagic Fever (CCHF) is endemic in some areas. However no human cases of CCHF have ever been reported in Mongolia despite the severe nature of the disease, and no investigation of possible animal infection has ever been conducted.

Methods: High risk and low risk areas of Mongolia for the presence of CCHF were identified by risk mapping using evidence from the scientific literature. This was undertaken by producing both a habitat suitability map for ticks and a risk map for human CCHF in HandiMap, a disease risk mapping program. A cross-sectional serological investigation for the presence of antibodies to CCHF virus among transhumant herders and their sheep was implemented in selected districts of Mongolia that were predicted to be high and low risk for the occurrence of CCHF virus respectively.

Results: Two types of risk maps were developed for Mongolia. One predicted the spatial distribution of the likelihood that the main tick vector of CCHF virus, *Hyalomma spp.* is present at a specific location, while the other predicted the likelihood that people in a particular area had been previously exposed to CCHF virus. Antibodies to CCHFV were detected in either sheep or human blood in all 21 districts sampled in the high risk area in the south of Mongolia, and in both species in 15 of 21 districts. One district in the Province of Selenge in the low risk area, was confirmed to be negative for anti-CCHFV IgG in both sheep and human blood samples.

Conclusion: HandiMap is a useful tool in facilitating disease risk assessment and guiding an epidemiological study to detect evidence of infection or demonstrate freedom from a disease. The study has confirmed for the first time the presence of CCHF in areas of Mongolia predicted to be at risk of infection. However, further studies are needed to determine why no human cases have been reported, and to better understand the ecology of hosts and epidemiology of the disease.

159

6.2. Introduction

Crimean Congo Hemorrhagic Fever is an acute highly contagious viral zoonosis caused by a member of the genus Nairovirus, family Bunyaviridae (Whitehouse, 2004). It is mainly transmitted to human by ixodid tick bites, or by direct contact with blood or tissues of viraemic hosts, such as during slaughter and butchering of ruminants. Hosts include cattle, sheep and goats, a range of wild animals, and some birds. Animal infection is subclinical. In most cases, human infections with CCHF virus are asymptomatic. However it can cause severe illness, with a fatality rate ranging from 2% to 80% (Çevik, Erbay et al., 2007; Chinikar, Goya et al., 2008; Ergönül, Çelikbaş et al., 2004; Hoogstraal, 1979; Vorou, Pierroutsakos, & Maltezou, 2007; Whitehouse, 2004). CCHF can also cause community and nosocomial outbreaks (Altaf, Luby et al., 1998; Ergönül, Çelikbaş et al., 2004; Parlak, Ertürk et al., 2015; Van Eeden, Joubert, & Van de Wal, 1985).

CCHF is endemic in substantial parts of Africa and Eurasia. Human infection with CCHF virus has been increasingly reported in recent years, mainly because of anthropogenic factors (Hoogstraal, 1979; Jameson, Ramadani, & Medlock, 2012; Leblebicioglu, 2010; Messina, Pigott et al., 2015; Randolph & Rogers, 2007). It has been predicted that the disease would likely expand beyond its traditional geographic locations as a result of climate change (Estrada-Pena & Venzal, 2007; Randolph & Rogers, 2007). Because of its public health importance, CCHF has been listed as a disease requiring notification to WHO under the revised International Health Regulations (2005) for notification of health events to the WHO (Formenty, Schnepf et al., 2007; Maltezou, Papa et al., 2009).

One common denominator for active CCHFV transmission is the presence of *Hyalomma spp.* ticks, since these are the principal vectors. A recent risk analysis of CCHF in Mongolia conducted as part of this study suggested that most parts of Mongolia were suitable for *Hyalomma* tick presence and there existed a number of locations where *Hyalomma* ticks (and other ixodid species which can be secondary vectors) could complete their lifecycle in small and large mammals (Roger Hewson, 2013). In Asia, the disease has been reported among humans and livestock in neighboring China and other countries such as Russia, Tajikistan and Kazakhstan in close proximity to Mongolia (Atkinson, Chamberlain et al., 2013; Hoogstraal, 1979; Yashina, Petrova et al., 2003; Yen, Kong et al., 1985). Some parts of Mongolia, particularly the provinces in the south, have the same ecological typology and fauna as in the countries mentioned here. Besides, the tick infested area is likely to be expanded as the size of steppe, desert steppe and desert has been projected to increase and shift northward because of climate change (Angerer, Han, Fujisaki, & Havstad, 2008). Although serological and virological testing of CCHFV were conducted around 1987 in wild mammals, the results have not been published on any international peer reviewed journals hence their accuracy could not be assessed. There had been no documentation of the disease

presence in either domestic animals or human population in Mongolia before this study was conducted (Messina, Pigott et al., 2015).

A number of reasons prompted that investigation of CCHFV should be implemented in Mongolia. Firstly, livestock in Mongolia represents a crucial national resource that offers an essential basis for sustainable development and maintenance of the herders' livelihood tradition. Agriculture sector produced 33 percent of gross domestic product, of which 87 percent was from livestock production (Fermet, Jane, & Forman, 2007). More than 90 percent of people employed in agricultural sector were engaged in animal husbandry in 2010. The traditional transhumant animal raising practices expose herders to tick bites. If ticks are infected with CCHFV, human infection would be inevitable. Secondly, the growing mining industry would increase the opportunities for susceptible miners working in the field to be exposed to ticks and possible infection with CCHFV. In the Gobi Desert, which lies across the southern part of Mongolia, herder family encampments can be up to 120 km apart because the population is very sparse and herders have very restricted access to medical services. The disease might exist in the country undetected because of low awareness and medical staff may lack the necessary disease-specific expertise to diagnose the disease.

The study was designed and conducted in 2013. It aimed to use HandiMap, a risk mapping module, to assist a risk-based survey investigating the evidence on the possible existence of CCHFV infection among domestic animals and/or human population in the Country.

6.3. Methods and Data

6.3.1. Study area
Mongolia lies between 52°9' north, 41°38' north, 87°47' east and 119°53' east. The selection of the locations for field investigation was guided by risk analysis and risk map(s).

6.3.2. Investigation approach
A review of the scientific literature on ecological factors which influence the host ticks and CCHF was implemented as the first step of the study. The objective of the review was to identify putative environmental and climatological predictors for host tick distribution, tick abundance and human CCHF cases, for which suitable remote-sensed data was available (Ansari, Shahbaz et al., 2014; Estrada-Peña, 2005; Estrada-Peña, Jameson et al., 2012; Estrada-Peña, Vatansever et al., 2010; Estrada-Peña & Venzal, 2007; Estrada-Peña, Zatansever et al., 2007; Roger Hewson, 2013; Vescio, Busani et al., 2012; Wilson, Gonzalez et al., 1991). The results are summarized in Table 6-1. The findings from the literature review

161

were used to inform the development of CCHFV suitability maps in Mongolia by using HandiMap, the risk mapping module of HandiResponse Program. Guided by the risk maps developed, targeted field investigation among domestic animals and herders was implemented in possible geographical hot spots and cold spots, in order to identify whether there was any evidence of the presence of CCHFV infection in Mongolia.

Table 6-1. Putative environmental and climatological factors for CCHFV suitability mapping

Environmental descriptor	Tick distribution	Tick abundance	Incidence of human infection
Temperature			
Mean monthly temperature		+	+
Absolute maximum temperature	+	+	
Winter temperature previous year			-
NDVI			
Mean monthly NDVI	+	+	+
NDVI anomaly			+
% of land covered by shrub, grass and herbaceous vegetation		+	+
Fragmentation			
Traversability, recruitment index		+	+
Number of different landcover types			+
SD* of mean NDVI			+
Annual mean precipitation	-	-	-

Note: *standard deviation; "+" means increase in the value of the environmental/climatological descriptor is significantly associated with an increase in the possibility or incidence of either tick distribution, abundance or CCHF infection among human. "-"means increase in the value of the environmental/climatological descriptor is significantly associated with a decrease in tick abundance or CCHF infection among human. Blank means the factor is not significantly associated with the measure, or has not been investigated. NDVI stands for normalized difference vegetation Index.

6.3.3.Development of CCHFV suitability map

Habitat suitability map for tick distribution (named as CCHF I map) and risk map for CCHF occurrence (CCHF II map) were developed. Following the ecological niche approach and guided by the four principles of relevance, comprehensiveness, no redundancy and measurability for selection of risk layers for GIS mapping (Malczewski, 2000), we chose mean NDVI from April to August, monthly maximum land-surface temperature, and mean annual precipitation layers for the development of CCHF I map, and percent of land covered by shrub, grass and herbaceous vegetation, mean NDVI from April to August, mean annual precipitation and mean maximum land surface temperature from April to August for CCHF

II-1 map. Because in the Mongolian environment, sparse vegetation and land covered by shrub, grass and herbaceous vegetation are negatively correlated spatially, a second CCHF prediction map was prepared, using sparse vegetation. For development of this CCHF II-2 map, mean NDVI from April to August, mean annual precipitation, mean maximum land surface temperature from April to August and sparse vegetation layer were used (Table 6-2).

Table 6-2. Selected putative environmental/climatological predictors for CCHF mapping in Mongolia

Environmental/climatological descriptor	Mapping purpose	Data source
% of land covered by shrub, grass and herbaceous vegetation	CCHF	http://earthenv.org/
Sparse vegetation, grassland, shrub land	CCHF	http://neo.sci.gsfc.nasa.gov/
Mean NDVI from April to August	Tick distribution, CCHF	http://neo.sci.gsfc.nasa.gov/
Annual mean precipitation	Tick distribution, CCHF	http://worldclim.org/
Maximum land-surface temperature mean from April to August	Tick distribution, CCHF	http://neo.sci.gsfc.nasa.gov

Values of all selected environmental descriptors were standardized between 0 and 1. Assumed relationship between a given descriptor layer and CCHF are documented in Table 6-3.

Table 6-3. Assumed relationship between a selected environmental and climatological descriptor and tick distribution, human CCHF infection risk

Environmental/climatological descriptor	Hypothesized relationship
% of land covered by shrub, grass and herbaceous vegetation	The higher percentage covered by grass, the lower the risk
Sparse vegetation, grassland, shrub land	Sparseness level is positively associated with increased risk
Mean NDVI from April to August	the higher NDVI value, the higher risk
Annual mean precipitation	The higher precipitation level, the lower risk
Maximum land-surface temperature mean from April to August	The higher maximum temperature level, the higher risk

Analytical Hierarchy Process (AHP) was used for weight assignment for the selected environmental and climatological descriptors. The relative importance between the selected environmental descriptors was informed by the study done by Messina et al. (Messina, Pigott et al., 2015). The results of weight assignments are documented in Table 6-4.

163

Table 6-4. Weight assignment for selected environmental/climatological descriptors for CCHF mapping in Mongolia

Risk layers	Tick distribution	CCHF II-1	CCHF II-2
% of land covered by shrub, grass and herbaceous vegetation		0.65	
Sparse vegetation, grassland, shrub land			0.65
Mean maximum temperature from April to August	0.80	0.21	0.21
Mean annual precipitation	0.11	0.08	0.08
Mean NDVI from April to August	0.09	0.06	0.06
Consistency Ratio	0.01	0.02	0.02

Note: NDVI stands for normalized difference vegetation index. Consistency Ratio is a comparison between consistency index and random consistency index. In case CR is equal or less than 0.1, the inconsistency level is acceptable.

6.3.4. Field investigation

A cross sectional CCHF sero-prevalence survey was conducted among herders and their sheep in twenty two districts from six provinces in August, 2013. The study was approved by the ethical committees from both Ministry of Health and Ministry of Food, Agriculture and Light Industry. Informed consent was obtained from the participating herder families. A brief interview with the herders was implemented by the investigation teams with a pre-designed questionnaire.

In each selected district, on average 20 herders from different families were interviewed and whole blood sample were taken from them. Selection of herders' families was based upon logistic feasibility. About 80 sheep from the same twenty families per district were bled for blood samples during summer time in 2013. This will give at least 95% confidence to detect at least one positive sample for the test method with a sensitivity of 80% at the designed disease prevalence of 5% (Cannon & Roe, 1982).

The recombinant anti-CCHFV NP IgG ELISA was performed for blood samples collected from sheep and human, selected sheep blood samples were tested by using indirect immunofluorescence antibody (IFA) method (Kranzler, Davidovich et al., 2013; Morikawa, 2013).

6.3.5. Concordance analysis

Cohen's Kappa was used to assess whether the presence of one or more positive sheep (versus zero positives) was associated with the presence (or absence) of positives in people from the same district.

6.4. Results

CCHF I, CCHF II-1 and CCHF II-2 maps (Figure 6-1,2,3) all suggest the southern Mongolia districts are high risk area for CCHF and the host ticks, while northern areas are at low risk.
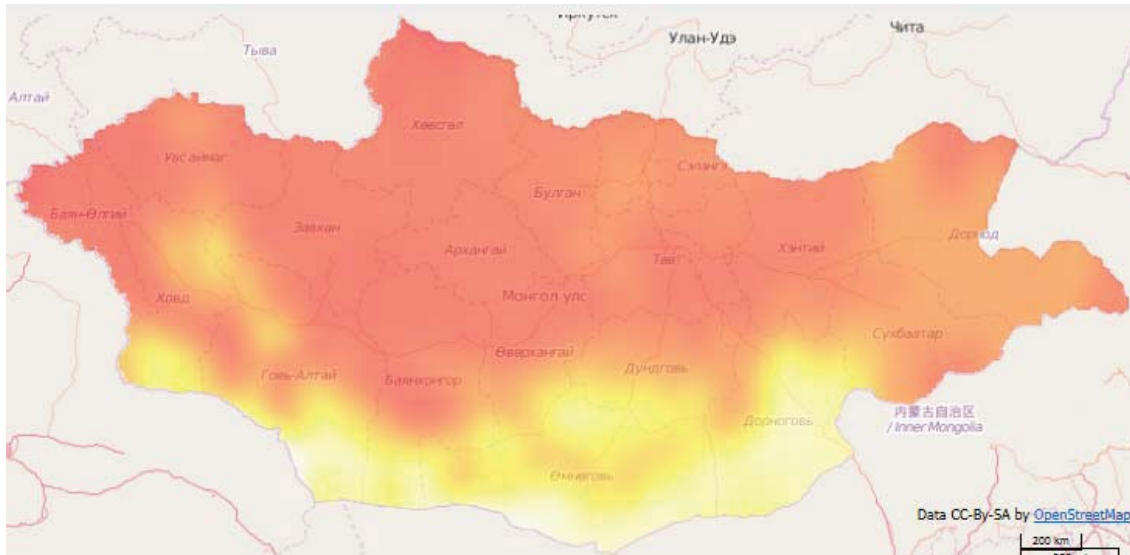
Figure 6-1. Habitat suitability map for tick distribution (CCHFV I map)
Note: the lighter the color the higher the probability for the presence of ticks.
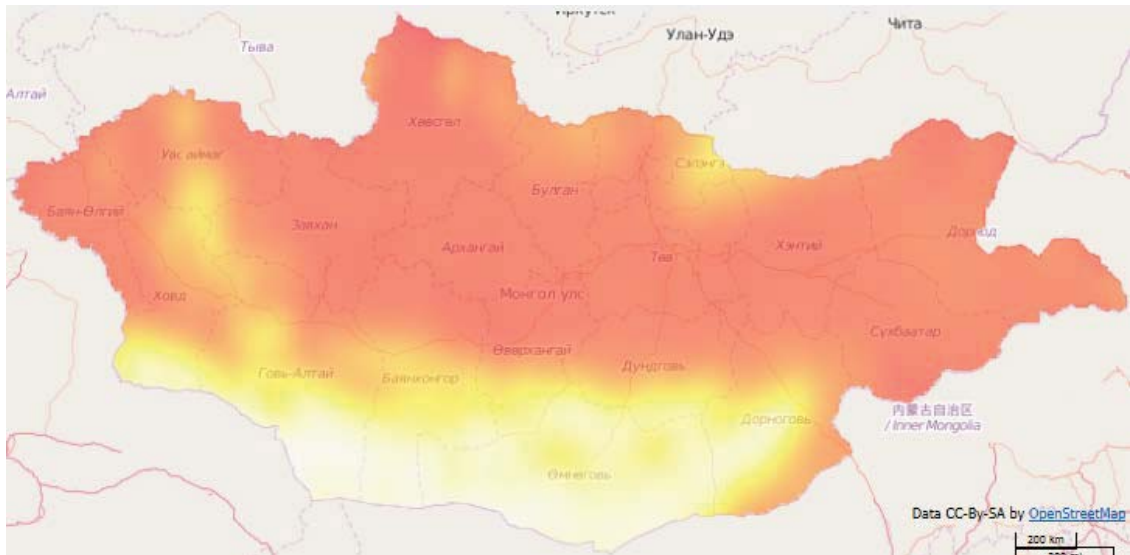
Figure 6-2. Risk map for CCHF occurrence (CCHFV II-1 map)
Note: the lighter the color the higher the probability for CCHF occurrence.

165
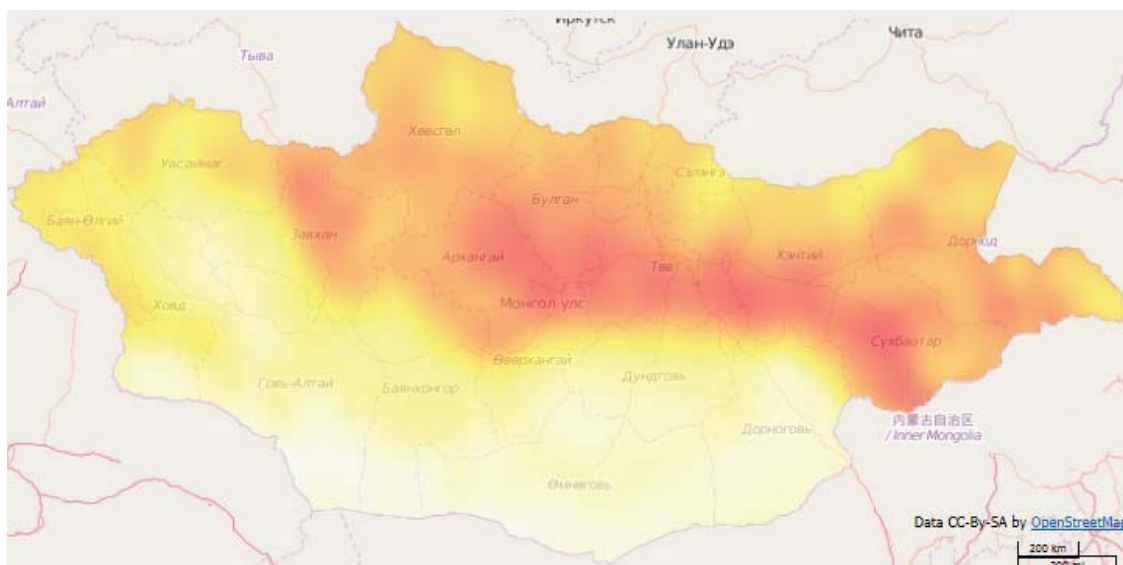
Figure 6-3. Risk map for CCHF occurrence (CCHFV II-2 map)
Note: the lighter the color the higher the probability for CCHF occurrence.

The number of sheep by district is illustrated in Figure 6-4 and the districts selected for field investigation are indicated in Figure 6-5. One southern district on the far right is very small, so not obvious on the map at this scale.

Range
☐ 0 - 48081
☐ 48081 - 96162
☐ 96162 - 144242
■ 144242 - 192123
■ 192123 - 240404
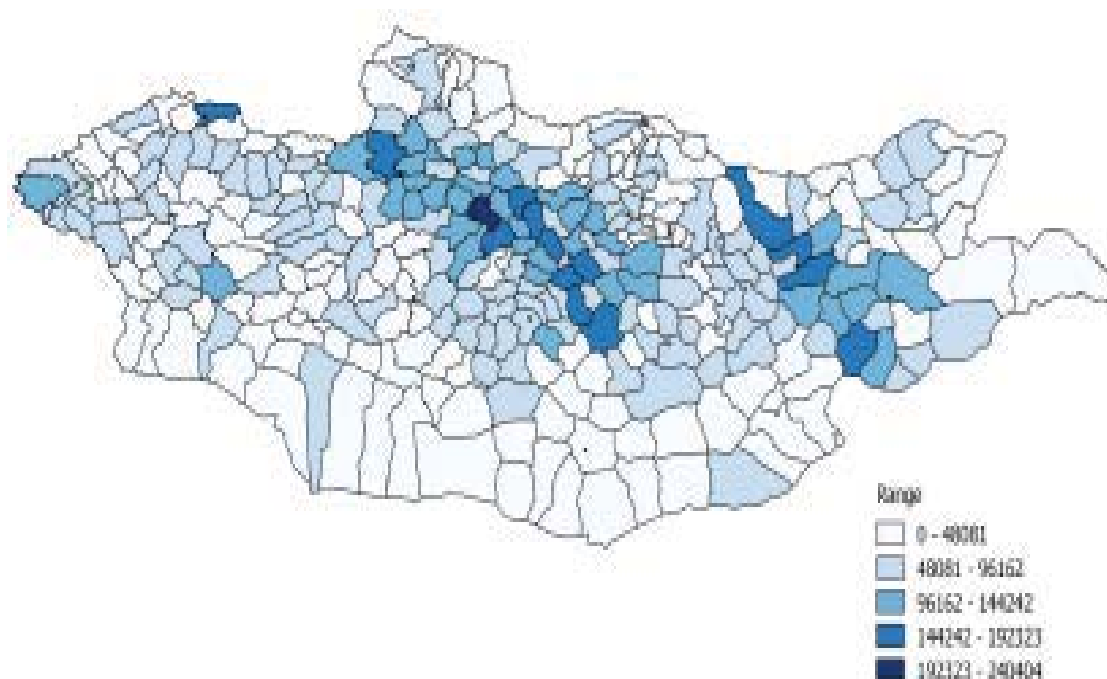
Figure 6-4. Number of sheep by district in Mongolia, 2013

166

Figure 6-5. Districts selected for field investigation of CCHF in Mongolia, 2013

The districts investigated were 21 high risk Gobi Desert districts along the border with China and 1 low risk northern district, as informed by the risk maps. The serological test results are summarized in Table 6-5.

Table 6-5. Anti-CCHFV IgG prevalence among human and sheep by district surveyed

Province	County	Sheep		Human	
		No. tested	% positive	No. tested	% positive
Khovd	Bulgan	80	17.5%	13	53.8%
	Yench	80	28.8%	22	36.4%
	Altai	80	36.3%	12	33.3%
Govi-Altai	Bugat	80	35.0%	24	20.8%
	Altai sum	80	42.5%	15	40.0%
	Tsogt	80	20.0%	13	15.4%
	Erdene	80	8.8%	13	7.7%
Bayankhongor	Bayan-Undur	80	6.3%	34	55.9%
	Shine jinst	80	21.3%	28	67.9%
Dornogovi	Erdene	75	0	26	15.4%
	Zamiin Ud	75	8.0%	23	8.7%
	Ulaanbadrakh	25	76.0%	26	11.5%
	Khuvsgul	75	4.0%	21	0
	Khatanbulag	75	0	21	57.1%
Umnugovi	Khanbogd	75	2.7%	23	30.4%
	Bayan-Ovoo	75	13.3%	37	62.2%
	Nomgon	75	0	20	60.0%
	Khurmen	75	24.0%	24	12.5%
	Bayandalai	75	9.3%	23	0
	Noyon	75	5.3%	22	0
	Gurvan	145	0.7%	26	30.8%
Selenge	Baruunburen	96	0	30	0

Note: confidence interval have not been provided since the study is intended to confirm whether a district has or has not serological evidence of CCHF.

Cohen's Kappa value is 0.083 and 95% confidence interval is -0.54 to 0.71. Therefore although only six of 22 districts are discordant, with the small number of total observations the results fail the concordance test. This means the result of whether a district is a CCHF positive among sheep is not correlated with whether it is a CCHF positive among human beings, and vice versa.

6.5. Discussion

6.5.1. CCHF infection exists in Mongolia

The study, for the first time, has generated serological evidence that CCHF infection has occurred in both people and domestic animals in Mongolia, and that the evidence of infection is present in areas of high ecological risk for both host ticks and the disease, but not in a low risk area. While this pattern of CCHF sero-positivity among both human and sheep in multiple locations suggests that the disease is likely to be endemic in at least some parts of the country, isolation and characterization of the virus would be required to provide definitive proof. The prevalence of anti-CCHFV IgG among both human and sheep is comparable to that in other CCHF endemic countries (Gonzalez, LeGuenno et al., 1990; Mostafavi, Haghdoost et al., 2013; Wilson, Gonzalez et al., 1991).

However, there has been no report of human CCHF cases in Mongolia so far. The reasons could include (i) only low pathogenic virus strain(s) exists in the country. There are seven known clades of CCHFV and they vary in virulence (Ergönül, 2012; Mild, Simon et al., 2010). Human CCHF cases have only been reported in locations where highly pathogenic CCHFV strains exist, such as Europe 1 (Hewson et al., 2004); (ii) Human cases of CCHF have occurred but clinically affected people did not seek health care or died before it could be obtained, and (iii) CCHF affected people sought health care but were misdiagnosed by health practitioners who were inadequately equipped with knowledge of the disease (Estrada-Peña, Jameson et al., 2012). Health personnel in Mongolia who were consulted were unaware of the possibility that this disease might be present, so a training program was arranged after the findings of this investigation were made available.

6.5.2. Selection of risk factors

Neither the human population density nor sheep density were included as an explanatory risk factor for CCHF prediction. Neither of them has been confirmed to be correlated with the incidence of reported human CCHF cases elsewhere (Messina, Pigott et al., 2015; Vescio, Busani et al., 2012). However reliance on reporting of CCHF cases only can be misleading since a significant proportion of CCHF infected persons could be asymptomatic or have mild clinical manifestation (Ergönül, Çelikbaş et al., 2004; Wilson, Gonzalez et al., 1991). Besides, we have not found any studies that either confirm or reject whether the density of human population or sheep is correlated to the presence of anti-CCHF IgG among these two populations. Hence, to be on the safe side, we decided not to use these population density risk layers. Density of both human and sheep populations are very low in the Gobi Desert, but the tick-transmitted nature of this disease make transmission possible even in areas of low density, especially

when families and their flocks are very mobile, as is the case with this transhumant (verging on nomadic) community.

For the development of human CCHFV infection, we attributed an inverse relationship between percentage of land covered by grass, shrub and and human CCHF occurrence. Such a relationship is different from that reported by some investigators (Messina, Pigott et al., 2015). A positive relationship was attributed between areas with sparse vegetation and CCHF risk. Areas with low percentage grass and shrub and sparse vegetation coverage mainly concentrate in southern Mongolia. Such an ecological environment in southern Mongolia provides suitable habitats for susceptible wild and domestic herbivore CCHFV hosts, and enables tick survival and virus circulation (Hewson, 2013). Besides, lower percentage coverage of grass and shrub also means high habitat fragmentation in Mongolia, which is also a risk factor for human CCHF infection (Estrada-Peña, Vatansever et al., 2010).

6.5.3. Selection of sheep for the CCHF survey

We chose sheep to be tested for anti-CCHFV IgG because (i) most of Mongolia herder families rear sheep and the number of sheep was almost four times higher than the cumulative total of cattle, camels and horses in 2003; (ii) previous studies revealed that sero-prevalence of CCHF usually tends to be higher in small ruminants such as sheep and goats than among large ruminants in the same CCHF affected areas (Mohamed, Said et al., 2007; Telmadarraiy Ghiasi et al., 2010; Williams, Al-Busaidy et al., 2000); (iii) presence of anti-CCHF IgG in sheep represents relatively recent active transmission of the disease because of their shorter lifespan compared with other other susceptible domestic animals (Wilson, Gonzalez et al., 1991); (iv) the test has better validation evidence for sheep than for other animal species.

6.5.4. Inconsistency in CCHF occurrence between people and animals

All high risk districts yielded positives in at least one species. However the test results from sheep samples did not always agree with the results from human samples on whether a high-risk district was CCHFV positive or negative. Among six districts that were in disagreement over CCHF status between people and sheep (Table 6-6), three districts were CCHF negative in sheep but positive in people and the remaining three were CCHF positive in sheep but negative in people.

Table 6-6. Districts in disagreement in CCHF status between human and sheep

District	Sheep	Human
Khuvsgul	4%	0
Bayandalai	9.3%	0
Noyon	5.3%	0
Erdene	0	26.7%
Khatanbulag	0	68.6%
Nomgon	0	38.7%

The first step in considering these different results is to confirm whether the laboratory results are largely valid.

Although false positive result cannot be ruled out for an individual test, this is highly unlikely for all the positive results. About twenty percent (341/1,640) of the sheep samples tested by ELISA were reassessed by IFA method, but we were not provided with data on the ELISA test status of samples submitted for IFA, other than that most but not all samples tested by IFA were ELISA positive. The results confirmed that quite a number of sheep samples were anti-CCHFV IgG positive by both methods. Even for the low prevalence districts like Gurvan in Umnugovi Province and Bayan-Under in Bayankhongor Province, these two methods were consistent in the investigation results (Morikawa, 2013).

Table 6-7. CCHFV antibody test results for sheep samples in selected districts by ELISA and IFA

Province	District	No. tested for ELISA	Positive (ELISA)	No. ELISA sample tested by IFA	Positive (IFA)
Khovd	Bulgan	80	17.5%	18	83.3%
	Yench	80	28.8%	30	96.7%
	Altai	80	36.3%	38	71.3%
Govi-Altai	Bugat	80	35.0%	34	55.9%
	Altai sum	80	42.5%	38	44.7%
Bayankhongor	Bayan-Undur	80	6.3%	2	50.0%
	Shine jinst	80	21.3%	4	75.0%
Umnugovi	Gurvan	145	0.7%	81	1.2%
Selenge	Baruunburen	96	0%	96	0.0%

Note: both ELISA and IFA tests were implemented at the Department of Veterinary Science, National Institute of Infectious Disease Japan.

False negative results of anti-CCHFV IgG could also happen in individual cases, but not to all infected sheep or herders in a district. The ELISA method for detection antibody to CCHFV is highly sensitive (Burt, Swanepoel, & Braack, 1993; Morikawa, 2013; Vanhomwegen, Alves et al., 2012). Newly infected animals or people can be seronegative, as can cases in very late convalescence and some very severe

human cases (Ergönül, 2012; Ergönül & Whitehouse, 2007; Papa, Mirazimi et al., 2015). None of these issues explain the discrepancy between sheep and human results in the six districts.

The issue may simply be a matter of sample size, particularly with respect to people, since the sample sizes for people would not reliably detect prevalences at district level below about 15%, depending on the district. Another explanation is that multiple transmission routes are occurring, and not all of them operate in all districts. People can get directly infected from tick bites, and the ticks may have become infected from wild animal hosts, rather than sheep. In some districts, depending on the ecological situation, infection may only be present in wild animals. People may also get infected from such direct contact tasks as slaughtering and/or butchering wild animals rather than sheep, in some cases. People may also become infected from other people. Sheep may be infected but there may be no transmission to people, because the particular people who were sampled do not adopt risky practices. Local tick populations may be small, and people may therefore escape exposure. In 15 districts both species are infected, and only in three districts are all sampled people negative, and in the other three districts all sampled sheep are negative.

6.5.5. Timing for field investigation

The investigation was implemented in August since this month was likely to be the peak season for infection among domestic animals and human as suggested by the epidemiological intelligence from the Xinjiang Uygur Autonomous Region in neighboring China (Saijo, 2007).

6.5.6. Interpretation of the risk landscapes

The risk landscapes prepared for this study using information from the scientific literature successfully identified large areas where CCHF was likely to establish if it entered Mongolia, and an area where it was unlikely to be present, and surveillance conducted in these areas provided strong evidence for the first time that CCHF virus is present in Mongolia, which was unexpected for national government personnel in human and animal health.

Exactly as found in this case, risk landscapes provide guidance on priority areas for surveillance, but do not provide any more information than that. The three maps which were produced differ in the size of the area which they indicate ranks towards the high end of the risk scale, because they use different risk factors, but if they were all overlaid then the summary map would show the highest risk area as the part of Mongolia where positive serology was found, and the negative area would fall in the low risk part. To further test which of the risk landscape maps is most accurate, it would be necessary to sample people and animals more widely across Mongolia, and this would clarify the predictive value of each map more fully.

172

6.5.7. Limitation of the study and future directions

The study presents only an initial step in understanding the complexity of CCHF in Mongolia. We hope it can prompt the Government to adopt a one health approach for further study on the disease, and develop a risk-based surveillance and disease mitigation program. Further actions we would like to suggest are (i) the samples collected from sheep and herders need to be tested for Ig M antibody to CCHFV, which can help identify any early infections among the two populations; (ii) ticks need to be collected and tested for CCHFV presence. Globally, CCHFV has been isolated from thirty species from seven genera of ticks that were naturally infected (Estrada-Peña & Jongejan, 1999; Turell, 2007).

Field investigation will be necessary to test presence of CCHFV via PCR in different species of ticks as well as their host wild animals. Such investigation would improve the understanding of the ecology of CCHFV in Mongolia (Camicas, Wilson et al., 1991). It can validate and improve the CCHFV I map and detect new disease foci; (iii) virus needs to be isolated and its genome sequenced to understand its virulence and pathogenicity; (iv) epidemiological studies need to be implemented to understand risk factors for human exposure to the virus in Mongolia; (v) risk-based surveillance adopting one health concept is needed among wild life, domestic animals, herders and workers in abattoirs in at least high risk areas identified by the CCHF maps and (vi) based upon the findings from the studies and surveillance activities, interventions may need to be implemented in the country (Mertens, Schmidt et al., 2013).

6.6. Conclusion

CCHF is known to occur in countries surrounding Mongolia, but has not been detected in Mongolia and was thought to be absent. The disease risk maps generated in HandiResponse using published scientific evidence facilitated the identification of high and low risk areas for CCHF and its host ticks. When survey was conducted in both types of areas, serological evidence of infection was found in people and their associated sheep flocks in the highest risk part of the country, but not in a low risk area. The study provided evidence that CCHFV is endemic in Mongolia, which needs to be further strengthened by virological investigations using PCR and possibly virus isolation. Further actions are warranted to understand the ecology of the disease and its pathogenicity by applying a one health approach and undertaking further surveillance in people, domestic animal hosts and wildlife hosts, leading if appropriate to mitigation measures designed to counteract possible severe human cases of the disease.

6.7. Reference

Altaf, A., S. Luby, et al. (1998). "Outbreak of Crimean-Congo haemorrhagic fever in Quetta, Pakistan: contact tracing and risk assessment." Tropical Medicine & International Health **3**(11): 878-882.

Angerer, J., G. Han, et al. (2008). "Climate change and ecosystems of Asia with emphasis on Inner Mongolia and Mongolia." Rangelands **30**(3): 46-51.

Ansari, H., B. Shahbaz, et al. (2014). "Crimean-Congo hemorrhagic fever and its relationship with climate factors in southeast Iran: a 13-year experience." The Journal of Infection in Developing Countries **8**(06): 749-757.

Atkinson, B., J. Chamberlain, et al. (2013). "Identification and analysis of Crimean-Congo hemorrhagic fever virus from human sera in Tajikistan." International Journal of Infectious Diseases **17**(11): e1031-e1037.

Burt, F., R. Swanepoel, et al. (1993). "Enzyme-linked immunosorbent assays for the detection of antibody to Crimean-Congo haemorrhagic fever virus in the sera of livestock and wild vertebrates." Epidemiology and infection **111**(03): 547-558.

Camicas, J.-L., M. Wilson, et al. (1991). Ecology of ticks as potential vectors of Crimean-Congo hemorrhagic fever virus in Senegal: epidemiological implications. Hemorrhagic fever with renal syndrome, tick-and mosquito-borne viruses, Springer**:** 303-322.

Cannon, R. and R. T. Roe (1982). Livestock disease surveys: a field manual for veterinarians, Australian Government Publishing Service.

Çevik, M. A., A. Erbay, et al. (2007). "Viral load as a predictor of outcome in Crimean-Congo hemorrhagic fever." Clinical infectious diseases **45**(7): e96-e100.

Chinikar, S., M. Goya, et al. (2008). "Surveillance and Laboratory Detection System of Crimean-Congo Haemorrhagic Fever in Iran." Transboundary and emerging diseases **55**(5-6): 200-204.

Ergönül, O. (2012). "Crimean–Congo hemorrhagic fever virus: new outbreaks, new discoveries." Current opinion in virology **2**(2): 215-220.

Ergönül, Ö., A. Çelikbaş, et al. (2004). "Characteristics of patients with Crimean-Congo hemorrhagic fever in a recent outbreak in Turkey and impact of oral ribavirin therapy." Clinical infectious diseases **39**(2): 284-287.

Ergönül, O. and C. Whitehouse (2007). "Crimean-Congo Hemorrhagic Fever, A Global Perspective." Springer. ISBN **10**: 1402061056.

Estrada-Peña, A. (2005). "Effects of habitat suitability and landscape patterns on tick (Acarina) metapopulation processes." Landscape ecology **20**(5): 529-541.

Estrada-Peña, A., L. Jameson, et al. (2012). "Unraveling the ecological complexities of tick-associated Crimean-Congo hemorrhagic fever virus transmission: a gap analysis for the western Palearctic." Vector-Borne and Zoonotic Diseases **12**(9): 743-752.

Estrada-Peña, A. and F. Jongejan (1999). "Ticks feeding on humans: a review of records on human-biting Ixodoidea with special reference to pathogen transmission." Experimental & applied acarology **23**(9): 685-715.

Estrada-Peña, A., Z. Vatansever, et al. (2010). "The trend towards habitat fragmentation is the key factor driving the spread of Crimean-Congo haemorrhagic fever." Epidemiology and Infection **138**(08): 1194-1203.

Estrada-Peña, A. and J. M. Venzal (2007). "Climate niches of tick species in the Mediterranean region: modelling of occurrence data, distributional constraints, and impact of climate change." Journal of Medical Entomology **44**(6): 1130-1138.

Estrada-Peña, A., Z. Zatansever, et al. (2007). "Modelling the spatial distribution of Crimean-Congo hemorrhagic fever outbreaks in Turkey." Vector-Borne and Zoonotic Diseases **7**(4): 667-678.

Fermet, E., R. Jane, et al. (2007). "Performance, Vision and Strategy." A tool for the governance of Veterinary Services **5**.

Formenty, P., G. Schnepf, et al. (2007). International surveillance and control of Crimean-Congo hemorrhagic fever outbreaks. Crimean-Congo Hemorrhagic Fever, Springer**:** 295-303.

Gonzalez, J.-P., B. LeGuenno, et al. (1990). "A fatal case of Crimean-Congo haemorrhagic fever in Mauritania: virological and serological evidence suggesting epidemic transmission." Transactions of The Royal Society of Tropical Medicine and Hygiene **84**(4): 573-576.

Hewson, R. (2013). Risk analysis of Crimean Congo Haemorrhagic Fever in Mongolia & Surveillance planning. Washington DC, USA.

Hewson, R., J. Chamberlain, et al. (2004). "Crimean-Congo haemorrhagic fever virus: sequence analysis of the small RNA segments from a collection of viruses world wide." Virus research **102**(2): 185-189.

Hoogstraal, H. (1979). "Review Article: The Epidemiology of Tick-Borne Crimean-Congo Hemorrhagic Fever in Asia, Europe, and Africa." Journal of medical entomology **15**(4): 307-417.

Jameson, L. J., N. Ramadani, et al. (2012). "Possible drivers of Crimean-Congo hemorrhagic fever virus transmission in Kosova." Vector-Borne and Zoonotic Diseases **12**(9): 753-757.

Kranzler, Y., N. Davidovich, et al. (2013). "A health in all policies approach to promote active, healthy lifestyle in Israel." Israel Journal of Health Policy Research **2**(1): 16-16.

Leblebicioglu, H. (2010). "Crimean–Congo haemorrhagic fever in Eurasia." International journal of antimicrobial agents **36**: S43-S46.

Malczewski, J. (2000). "On the use of weighted linear combination method in GIS: common and best practice approaches." Transactions in GIS **4**(1): 5-22.

Maltezou, H. C., A. Papa, et al. (2009). "Crimean-Congo hemorrhagic fever in Greece: a public health perspective." International Journal of Infectious Diseases **13**(6): 713-716.

Mertens, M., K. Schmidt, et al. (2013). "The impact of Crimean-Congo hemorrhagic fever virus on public health." Antiviral research **98**(2): 248-260.

Messina, J. P., D. M. Pigott, et al. (2015). "The global distribution of Crimean-Congo hemorrhagic fever." Transactions of The Royal Society of Tropical Medicine and Hygiene: trv050.

Mild, M., M. Simon, et al. (2010). "Towards an understanding of the migration of Crimean–Congo hemorrhagic fever virus." Journal of General Virology **91**(1): 199-207.

Mohamed, M., A.-R. Said, et al. (2007). "A serological survey of Crimean-Congo haemorrhagic fever in animals in the Sharkia Governorate of Egypt." Veterinaria italiana **44**(3): 513-517.

Morikawa, S. (2013). Serological Surveillance of Crimean Congo Haemorrhagic Fever Virus infection among Sheep in Mongolia. Japan, Department of Veterinary Science,National Institute of Infectious Diseases, Japan.

Mostafavi, E., A. Haghdoost, et al. (2013). "Spatial analysis of Crimean Congo hemorrhagic fever in Iran." The American journal of tropical medicine and hygiene **89**(6): 1135-1141.

Papa, A., A. Mirazimi, et al. (2015). "Recent advances in research on Crimean-Congo hemorrhagic fever." Journal of Clinical Virology **64**: 137-143.

Parlak, E., A. Ertürk, et al. (2015). "A nosocomial outbreak of Crimean-Congo hemorrhagic fever." Journal of Microbiology and Infectious Diseases **5**(01).

Randolph, S. E. and D. J. Rogers (2007). Ecology of tick-borne disease and the role of climate. Crimean-Congo hemorrhagic fever, Springer**:** 167-186.

Saijo, M. (2007). Crimean-Congo hemorrhagic fever in the Xinjiang Uygur autonomous region of Western China. Crimean-Congo Hemorrhagic Fever, Springer**:** 115-130.

Telmadarraiy, Z., S. M. Ghiasi, et al. (2010). "A survey of Crimean-Congo haemorrhagic fever in livestock and ticks in Ardabil Province, Iran during 2004-2005." Scandinavian journal of infectious diseases **42**(2): 137-141.

Turell, M. J. (2007). Role of ticks in the transmission of Crimean-Congo hemorrhagic fever virus. Crimean-Congo Hemorrhagic Fever, Springer**:** 143-154.

Van Eeden, P., J. Joubert, et al. (1985). "A nosocomial outbreak of Crimean-Congo haemorrhagic fever at Tygerberg Hospital. Part I. Clinical features."

Vanhomwegen, J., M. J. Alves, et al. (2012). "Diagnostic assays for Crimean-Congo hemorrhagic fever." Emerging infectious diseases **18**(12): 1958.

Vescio, F. M., L. Busani, et al. (2012). "Environmental correlates of crimean-congo haemorrhagic fever incidence in Bulgaria." BMC public health **12**(1): 1116.

Vorou, R., I. N. Pierroutsakos, et al. (2007). "Crimean-Congo hemorrhagic fever." Current opinion in infectious diseases **20**(5): 495-500.

Whitehouse, C. A. (2004). "Crimean–Congo hemorrhagic fever." Antiviral research **64**(3): 145-160.

Williams, R., S. Al-Busaidy, et al. (2000). "Crimean-Congo haemorrhagic fever: a seroepidemiological and tick survey in the Sultanate of Oman." Tropical Medicine & International Health **5**(2): 99-106.

Wilson, M., J.-P. Gonzalez, et al. (1991). Epidemiology of Crimean-Congo hemorrhagic fever in Senegal: temporal and spatial patterns. Hemorrhagic Fever with Renal Syndrome, Tick-and Mosquito-Borne Viruses, Springer**:** 323-340.

Yashina, L., I. Petrova, et al. (2003). "Genetic variability of Crimean-Congo haemorrhagic fever virus in Russia and Central Asia." Journal of General Virology **84**(5): 1199-1206.

Yen, Y., L. Kong, et al. (1985). "Characteristics of Crimean-Congo hemorrhagic fever virus (Xinjiang strain) in China." The American journal of tropical medicine and hygiene **34**(6): 1179-1182.

7. Modelling alternative surveillance methods to detect incursion of an exotic zoonotic disease

7.1. Abstract

Background: Emerging diseases pose a constant threat to countries throughout the world, and an incursion of a novel disease could result in social and economic disruption in affected countries, as has occurred with avian influenza and Ebola virus disease. Timely detection of disease outbreaks in at-risk countries is critical. An objective approach is needed for predicting the epidemiological performance of various surveillance methods in order to identify optimum combinations of methods for use under the circumstances of an individual country.

Methods: We used a hypothesized zoonotic disease, Austeria, as an example to test the effects of different surveillance options in disease detection. A temporal and spatial dynamic model was developed for simulating the disease in Queensland Australia over a period of 365 days. Eight surveillance components were designed to detect the disease by focusing on different strata of affected pig or human populations. Each component has nine variant sub-components that differ in surveillance intensity and sensitivity of detection. Using the Austeria model, the efficiency of each surveillance method in detecting a disease incursion was evaluated, as measured by (i) number of days for outbreak detection (efficiency), (ii) proportion of simulated outbreaks detected (effectiveness), and (iii) number of farms infected by the date of detection (which strongly influences economic consequences of the outbreak). The evaluation was conducted by simulating 99 epidemiologically distinct outbreaks, then separately evaluating the performance of each surveillance component and its subcomponents on these 99 outbreaks.

Results: The most efficient approach was risk-based sampling of commercial herds in areas where there was substantial numbers of both commercial herds and feral pigs (strategy RB2), and the least efficient method was use of hunters to collect blood samples from feral pigs they captured (strategy FPB). The remaining six surveillance components fell in between and were clustered in the middle of the range between the two extreme results on each of the performance measures. Strategy RB2 with high surveillance intensity and high sensitivity of detection could detect the outbreak at a median of 108 days

after incursion, and successfully detected over 93 percent of the 99 simulated outbreaks. When this component and subcomponent was used, it would detect the disease when a median of 23 farms were infected. In contrast, the FPB component with low intensity and low sensitivity sub-components detected only 3 percent of the 99 simulated outbreaks, within this 3% it could only detect the disease at a median number of 281 days after the incursion, and when the disease was detected, a median of 2,515 farms were already infected.

Conclusions: The methodology developed for assessing performance of surveillance strategies allowed 72 different surveillance methods to be compared, using a combination of three evaluation criteria. The approach can be generalized for other infectious diseases.

7.2. Introduction

A recent estimate showed that infectious disease accounted for more than one third of the total global disease burden, and the share was even higher in developing countries (Murray, Vos et al., 2013). Among the infectious diseases, emerging zoonotic diseases have taken centre stage in recent years because of their potential for generating social, economic and humanitarian disruption and the growing numbers of outbreaks caused by these disease agents (Brahmbhatt, 2006; Fan, 2003; Meltzer, Cox, & Fukuda, 1999; Narrod, Zinsstag, & Tiongco, 2012; UNDG, 2015; WorldBank, 2014). Over the past four decades, we have seen on average one to three newly emerged infectious human diseases per year, and approximately 75 percent of the newly emerged human diseases are zoonotic (Jones, Patel et al., 2008; Taylor, Latham, & Mark, 2001). Globalization, international trade, increasing human population, habitat encroachment, changing agricultural practices have been either hypothesized or confirmed to contribute to such an increase (Barrett, Kuzawa et al., 1998; Bengis, Leighton et al., 2004; Brown, 2004; Daszak, Cunningham, & Hyatt, 2000; Jones, Patel et al., 2008; Patz, Daszak et al., 2004; Taylor, Latham, & Mark, 2001).

Surveillance is critical for responding to all infectious disease threats. Earlier detection would win more time to respond and hence potentially less infections will result (Kaufmann, Meltzer, & Schmid, 1997; Longini, Nizam et al., 2005). Various approaches have been proposed or used for improving timeliness of surveillance systems. These include: (i) improving quality of existing signals (for instance, by adopting active disease reporting or implementing risk-based surveillance); (ii) adding new signals (for instance implementation of different surveillance components and activities); (iii) improving detection algorithm and (iv) optimizing the detection strategy (Box, Jenkins, & Reinsel, 2013; Goutard, Paul et al., 2012; Jajosky & Groseclose, 2004; Martinez, 2000; Melton & Hripcsak, 2005; Wagner, Tsui et al., 2001). Surveillance approaches not only differ in technical performance and cost, but also in operational

complexity. In the real world, people tend to use a combination of different surveillance approaches or components in the expectation that this will improve detection. However, selection of an appropriate combination of surveillance methods has largely been based on subjective judgment rather than objective assessment.

This paper introduces a systematic approach to assessing the expected outcomes of various combinations of surveillance approaches for an infectious disease affecting both animals and human beings. The outcomes are assessed in terms of time to detection and economic effects of delayed detection (the latter will be presented in Chapter 8). It was conducted by using stochastic modelling of the disease and economic analysis of the costs and benefits of different combinations of surveillance methods. The goal is to determine an optimum mix of surveillance methods which in combination will be more effective, efficient, and/or of lower cost than using a single method. Most emerging diseases involve wild animals as well as domestic animals, so the assessment needs to consider surveillance in both wild and domestic animals, and in humans.

7.3. Methods and data

7.3.1. The example disease - Austeria

Austeria is a hypothesized infectious zoonotic disease, invented to demonstrate the evaluation of surveillance portfolios for detection of diseases involving domestic animals, wild or feral animals and people. It is caused by a virus affecting mainly pigs, and is modelled on the epidemiology of porcine reproductive and respiratory syndrome (PRRS) virus, although various features have been adjusted to create the purely artificial disease Austeria, including its ability to infect people. The virus can be transmitted via direct contact between an infected pig and a susceptible one, by fomites, or by aerosol transmission over a distance of at least 1 km. Pigs are susceptible to infection by both oral and respiratory routes. The muscles of infected pigs contain high concentration of virus. Therefore a pig may also be infected by ingesting feed contaminated with the virus.

The State of Queensland in Australia was chosen as the test location for the investigation because it has an appropriate mix of large and small scale owned pig herds, and a substantial feral pig population overlapping owned pig herds in spatial distribution. These populations are all adequately defined in spatial and demographic terms (including enterprise-level economic performance) to allow spatial modelling and economic analysis to be undertaken.

We assumed for this analysis that the disease was exotic to Australia and it was brought into the country by illegal introduction of pork contaminated by the virulent virus. The pork was subsequently discarded and included in food waste illegally fed to the primary case – a small "backyard" (non-commercial) pig herd, located in an area of Queensland where a substantial population of feral pigs is present, and there is opportunity for interaction between the pigs in the primary case herd and feral pigs. The owner of the primary case herd also bought and sold pigs to other small herds. The disease was initially transmitted to other backyard herds and then to commercial pig farms and feral pig families. Both local spread and multiple movement types contribute to the disease diffusion.

We also assumed the disease could cause human infection by direct contact between a person and an infected pig, but not by consumption of infected pork products. However, the disease could not generate human to human transmission.

7.3.2. Data needed for modelling

A spatial and temporal model predicting Austeria disease spread was developed by using HandiSpread software[5], the enhanced version of InterSpread Plus. The software was populated with estimated or actual locations of all commercial and non-commercial pig farms in Queensland.

One back yard pig herd located in an area of moderate density of feral pigs was selected as the starting point of the disease introduction for an outbreak simulation in Queensland. Ninety-nine replicates of each simulation were run using different random number seeds in order to establish a distribution of possible outcomes for an outbreak scenario; each outbreak was allowed to continue for a 365 day period without any controls. HandiSpread uses independent random number seeding of each process within the model, and so it is possible to generate exactly the same set of 99 different disease outbreaks that reflect a range of natural epidemiological variability as many times as needed. A range of surveillance strategies can then be modelled for the same set of 99 outbreaks and their performance summarized for each. The choice of 99 outbreaks was based on preliminary assessment of variability between replicates, and the use of an odd number of replicates makes selection of a "median outbreak" possible, and this median outbreak will be used extensively to demonstrate spatial results of model runs.

---

[5] InterSpread Plus; Massey University EpiCentre, Palmerston North, New Zealand. Available at http://www.interspreadplus.com, accessed August 22, 2015, 2015

7.3.2.1.   Number and location of domestic pigs and feral pigs

Estimates on size of individual domesticated pig herd, feral pig family and their locations were done by Neumann, et al[6].

Commercial pig herds

Data on the postcode area in which each registered pig herd that routinely sends pigs to an abattoir for slaughter (and is therefore considered to be commercial) was obtained from the PigPass database held by Australian Pork Limited. An official copy of the shape file describing the polygon outline of each Australian postcode was obtained[7] in order to assign a specific point location to each commercial pig farm. This point location was based on post code information attached to each farm in the PigPass database[8] which was provided by Australian Pork Limited. To ensure the confidentiality of individual property details, geographic information system software[9] was used to randomly assign a fixed location to each farm listed in PigPass, within its respective postcode area. Data on the pig industry in Queensland Australia are summarized in Table 7-1.

Table 7-1. Number of feral pig and commercial pig herds in Queensland, Australia, 2013

Type	Farm description	Sow (or boar) inventory	Growing pig inventory	Number of farms
Sm	Small registered commercial breeder farms (farrow to finish)	pert(1,3,7)	pert(0,2,7)	318
By	Backyard farms unregistered (farrow to finish)	pert(1,3,7)	pert(0,2,7)	1532
grow	Commercial growers and multi-site growers only (no breeders)	NA	pert(240,650,12000)	164
Lg	Large commercial breeders (farrow to finish)	pert(8,125,500)	pert(7,8,9)	196
multi-ff	Multi-site commercial farms (farrow to finish)	pert(200,350,3000)	pert(3,4,5)	21
multi-sow	Multi-site commercial breeders only	pert(200,350,3000)	NA	24
Ai	Artificial insemination	pert(10,20,30)	NA	2

[6] Neumann EJ, Hall WF, Morris RS, O'Leary B. The risk and consequences of PRRS virus introduction to Australia through importation of pork (Project 2011/1039.426). Australian Pork Ltd, Barton, ACT, Australia, 2013.
[7] Terrapages Pty Ltd. Australian Postcode spatial database (version dated August 2012); Pyrmont, NSW
[8] Australian Pork Limited. Fact Sheet: What is the PigPass system? http://www.pigpass.com.au/documents/FACTSHEET-WhatIsThePigPassSystem_Feb2011.pdf. Accessed November 18, 2012
[9] Quantum GIS Development Team (2012), version 1.8.0 Lisboa. Quantum GIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org

Back yard pig herds

No single register of non-commercial pig holdings exists in Australia and therefore information and data was accumulated from various official and unofficial sources to generate reasonable estimates of the total number of "backyard" pig holdings and their locations. The specific methods and sources are described by Neumann *et al* in the report mentioned above.

Estimates of the minimum, most likely, and maximum number of backyard pig farms (not listed in the PigPass database) was achieved through review of State Livestock Registers, previous published estimates and a survey among a group of experts. A simulation approach was used to generate the spatial distribution of the population of backyard pig herds, since in most cases there is no geographic data available for this group. The number of backyard pig herds was assumed to be in proportion to the density of the human population in each area, but limited to areas for which the principal land use is agriculture. For mesh blocks with population densities falling in the range of 0.1 to 5 persons per square kilometer, the probability of keeping backyard pigs was set to 0.1. For each selected mesh block a single random draw from the binomial distribution was used to estimate the number of households within the mesh block that kept backyards pigs. For each of these 'backyard pig households' a point location was assigned within the boundaries of the respective mesh block. The resulting spatial distribution was compared with those backyard pig farms having known geographic locations, and was considered as a credible representation of the distribution of backyard herds.

Feral pig families

The total number of feral pigs in Australia is quite uncertain, and various estimates have put the population at various levels ranging from 3.5 to 23 million (Hone, 1990), with exact numbers known to vary from year to year and between seasons. Feral pigs live in family groups, with home ranges depending on habitat suitability and availability of feed and water. Estimation of feral pig density was based on the relative suitability of the habitat for feral pigs (supplemented by count data where available) and was scaled from 1 ('not suitable') to 9 ('highly suitable') to enable the distribution of feral pig families to be mapped on to Queensland, to create a population of feral pig family units which could interact with owned pig herds in the model. The starting point for the development of the map was information provided by a group of experts in Australian vertebrates (Cowled, Giannini et al., 2009; Hone, 2012; West, 2008) and from this information a numerically scaled zoning map (Figure 7-1) of feral pig habitat suitability was created. No areas of Queensland were classified in zone 9.

For the purpose of this modelling, it was assumed that 10 million feral pigs were present in Australia, existing in family groups with a mean size of 10 pigs (adults, juveniles, and young), suggesting a total of

one million family groups in the country. These one million family groups were distributed across the country at densities proportional to the total land mass in each habitat suitability zone, and 429,262 of these families were located in Queensland. The breakdown of feral families by habitat zone in Queensland is shown in Table 7-2 below, and the distribution of owned pig herds by feral pig density zone is shown in

Table 7-3. Once the total number of feral family groups for each zone was determined, each family was randomly assigned a point location within its respective habitat zone. The estimated feral pig density distribution in Queensland is illustrated in Figure 7-1.

Table 7-2. Estimated distribution of feral families by risk zone in Queensland, Australia

Feral risk zone	Feral families	Area (km^2)	Area per feral family
1	600	13,686	22.8
2	152	1,923	12.7
3	1,341	9,952	7.4
4	102,253	573,655	5.6
5	915	3,936	4.3
6	203,820	765,431	3.8
7	2,132	7,225	3.4
8	118,049	347,634	2.9
Summary	**429,262**	**1,723,442**	**4.0**

Note: the higher number the feral risk zone, the more suitable for feral pigs.

Table 7-3. Number of owned herds of each type in the feral pig density zones

Density zone	1	2	3	4	5	6	7	8	Total
ai				2					2
by				592		837		103	1,532
feral	600	152	1,341	102,253	915	203,820	2,132	118,049	429,262
grow				93		66		5	164
lg				102		81		13	196
multi-ff				11		10			21
multi-sow				15		8		1	24
sm				200		113		5	318
Total	**600**	**152**	**1,341**	**103,268**	**915**	**204,935**	**2,132**	**118,176**	**431,519**

Note: refer to explanations on farm type in Table 7-1. Number in the first row represents suitability scale, ranges from 1 (poor suitability) to 9 (high suitability).
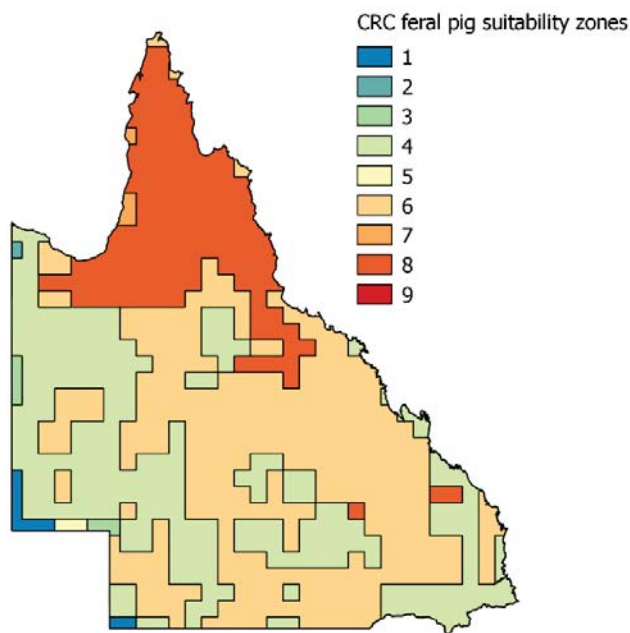
Figure 7-1. Estimated relative density of feral pigs in Queensland, Australia
Note: based on the suitability of habitat, the zoning was created using initial source data obtained from
the Invasive Animal Cooperative Research Centre (CRC), available at www.invasiveanimals.com (West,
2008). Suitability has been scaled from 1 (poor suitability) to 9 (high suitability).

7.3.2.2. Incubation period and infectivity

The mean incubation period for Austeria, based on 1000 draws from a log-normal distribution, was set at
8 days and with a standard deviation of three days. Infectivity of infected herds was set relative to the
onset of clinical signs with 60% infectiousness at Day 1 post-onset of clinical signs and 100%
infectiousness at Day 5 post-onset of clinical signs. Infectiousness then declined subsequently.

7.3.2.3. Infection transmission between herd types

Austeria is hypothesized to be spread by three types of local spread, four movements between different
sectors of the commercial and backyard herd populations, as well as interaction between feral pigs and
domestic pigs as depicted in Figure 7-2. The feral pig population is represented on the left hand side of
the diagram, and the various populations of owned pigs on the right. Local spread between pigs that are in
fairly close proximity but do not have a known action that transmitted infection. Movement spread is
relates to animals being moved deliberately or known forms of movement taking place.

Parameters such as frequency and distance for different types of movement, probability of transmission
for each type of movement were estimated after consultation with the experts and are documented in
Table 7-4 to Table 7-17.
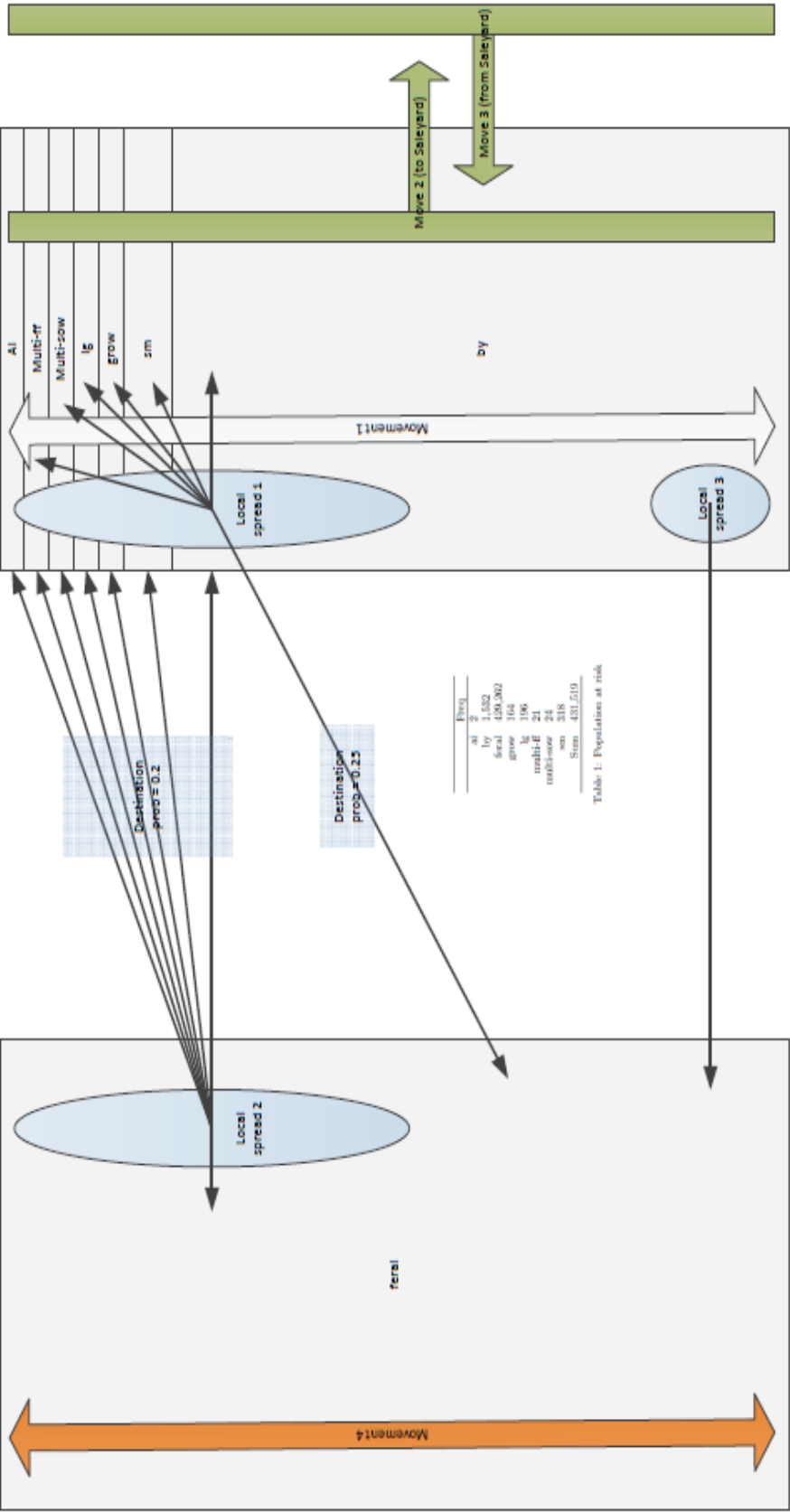
Figure 7-2. Infection transmission routes between herd types for pigs in Queensland, Australia

Table 7-4. Movement type 1: Simulates routine movements within and among the types of commercial and backyard pig holdings

Origin	Destination								Row sum	Frequency[a]
	lg	Sm	grow	ai	by	feral	multi-ff	multi-sow		
Lg	0.005	0.005	0.89	0	0.1	0	0	0	1.000	Poisson(0.2)
Sm	0	0.05	0.05	0	0.9	0	0	0	1.000	Poisson(0.1)
grow	0.1	0.1	0.1	0	0.5	0	0.1	0.1	1.000	Poisson(0.02)
Ai	0.3	0.1	0	0	0	0	0.3	0.3	1.000	Poisson(0.2614)
By	0	0.01	0	0	0.99	0	0	0	1.000	Poisson(0.055)
feral	0.01	0.01	0.04	0	0.92	0	0.01	0.01	1.000	Constant(0)
multi-ff	0.005	0.005	0.89	0	0.1	0	0	0	1.000	Poisson(0.2)
multi-sow	0.005	0.005	0.89	0	0.1	0	0	0	1.000	Poisson(0.2)

Note: feral pigs are excluded from these movements; [a] Daily movement frequency; for explanations on farm type refer to Table 7-1.

Table 7-5. Movement type 1: distance matrix

Proportion	0.5784	0.1671	0.0889	0.0546	0.034	0.024	0.0185	0.0145	0.0111
Distance (m)	50000	100000	150000	200000	250000	300000	350000	400000	450000

Probability of transmission = Constant (0.50)

Note: estimated by experts.

Table 7-6. Movement type 2: Simulates movement of commercial and backyard pig holdings to sale yards

| Origin | Destination | | | | | | | | Row sum | Frequency [a] |
	lg	sm	grow	ai	By	feral	multi-ff	multi-sow		
Lg	NA	NA	NA	NA	NA	NA	NA	NA	NA	Poisson(0.0174)
Sm	NA	NA	NA	NA	NA	NA	NA	NA	NA	Poisson(0.025)
grow	NA	NA	NA	NA	NA	NA	NA	NA	NA	Poisson(0.0016)
Ai	NA	NA	NA	NA	NA	NA	NA	NA	NA	Poisson(0.0)
By	NA	NA	NA	NA	NA	NA	NA	NA	NA	Poisson(0.025)
feral	NA	NA	NA	NA	NA	NA	NA	NA	NA	Constant(0)
multi-ff	NA	NA	NA	NA	NA	NA	NA	NA	NA	Poisson(0.0174)
multi-sow	NA	NA	NA	NA	NA	NA	NA	NA	NA	Poisson(0.0174)

Note: NA means not applicable; [a] Daily movement frequency; Destination is to a sale yard and the frequency of such movements from each herd type is determined by the various Poisson distributions in the column on the right. Pigs which go to a saleyard must then be transferred to the recipient herd, movements to abattoirs through saleyards are not considered, because they are not involved in disease spread. Onward transmission of infection from the saleyard to the recipient herd is shown in Table 7-8 and Table 7-9, and the probability of transmission during this second move is determined by the original source of the animals. In HandiSpread, the only movements which are represented in the modelling process are those where infection transmission is possible, movements of uninfected animals are not considered to represent a risk of transmission of infection and are not modelled.

Table 7-7. Movement type 2: movement distance matrix

Proportion	0.5784	0.1671	0.0889	0.0546	0.034	0.024	0.0185	0.0145	0.0111
Distance (m)	50000	100000	150000	200000	250000	300000	350000	400000	450000

Table 7-8. Movement type 3: simulates movement from sale yards to commercial and backyard pig holdings

Origin	Destination									Frequency [a]
	lg	Sm	grow	ai	by	feral	multi-ff	multi-sow	Row sum	
Lg	0.038	0.038	0	0	0.924	0	0	0	1.000	NA
Sm	0.038	0.038	0	0	0.924	0	0	0	1.000	NA
grow	0	0	0.2	0	0.8	0	0	0	1.000	NA
Ai	0	0.2	0	0	0.8	0	0	0	1.000	NA
by	0.005	0.005	0	0	0.99	0	0	0	1.000	NA
feral	0	0	0	0	0	1	0	0	1.000	NA
multi-ff	0.038	0.038	0	0	0.924	0	0	0	1.000	NA
multi-sow	0.038	0.038	0	0	0.924	0	0	0	1.000	NA

Note: NA means not applicable; [a] Daily movement frequency; Destination is from a sale yard.

Table 7-9. Movement type 3: distance matrix

Proportion	0.5784	0.1671	0.0889	0.0546	0.034	0.024	0.0185	0.0145	0.0111
Distance (m)	50000	100000	150000	200000	250000	300000	350000	400000	450000

Probability of transmission = Constant (0.50)

Table 7-10. Movement type 4: Long distance movements among feral pigs

| Origin | Destination | | | | | | | | Row sum | Frequency[a] |
	lg	sm	grow	ai	by	feral	multi-ff	multi-sow		
lg	0	0	0	0	0	1.000	0	0	1.000	Constant(0)
sm	0	0	0	0	0	1.000	0	0	1.000	Constant(0)
grow	0	0	0	0	0	1.000	0	0	1.000	Constant(0)
ai	0	0	0	0	0	1.000	0	0	1.000	Constant(0)
by	0	0	0	0	0	1.000	0	0	1.000	Constant(0)
feral	0	0	0	0	0	1.000	0	0	1.000	Poisson(0.08)
multi-ff	0	0	0	0	0	1.000	0	0	1.000	Constant(0)
multi-sow	0	0	0	0	0	1.000	0	0	1.000	Constant(0)

Note: [a] Daily movement frequency.

Table 7-11. Movement type 4: distance matrix

Proportion	0.1	0.6	0.29	0.01
Distance (m)	1000	5000	20000	50000

Probability of transmission = Constant (0.10)

Local spread type 1: Local spread of infection originating from all farm classes and spreading to all farm classes. Provides the only mechanism for infection to move from commercial farms to feral pigs (at a very low rate). Feral families are primarily infected through local spread from backyard pigs (Local spread type 3).

Table 7-12. Local spread 1: probability of transmission

Distance (m)	1000	2000	3000
Probability of transmission	0.012	0.003	0.001

Table 7-13. Adjustment in probability of transmission for specific farm class destinations

Farm class	Adjustment factor to reduce probability of transmission
lg	1
sm	1
grow	1
ai	0.0
by	1
multi-ff	1
multi-sow	1
feral	0.15

Local spread type 2: Local spread representing the only mechanism by which infection can move from feral pigs to commercial and backyard farms (low frequency event over short distance; origin of infection limited only to feral families).

Table 7-14. Local spread type 2: probability of transmission

Distance (m)	1000	2000	3000	5000
Probability of transmission	0.0045	0.002	0.0008	0.0005

Table 7-15. Local spread type 2: adjustment in probability of transmission for specific farm class destinations

Farm class	Adjustment factor to reduce probability of transmission
Lg	0.2
Sm	0.2
Grow	0.2
Ai	0.2
By	0.2
multi-ff	0.2
multi-sow	0.2
Feral	1

Local spread type 3: Local spread representing spread of infection from backyard pigs to feral (origin of infection limited only to backyard farm class).

Table 7-16. Local spread type 3: probability of transmission

Distance (m)	1000	2000	3000
Probability of transmission	0.0025	0.00125	0.001

Table 7-17. Local spread type 3: adjustment in probability of transmission for specific farm class destinations

Farm class	Adjustment factor to reduce probability of transmission
Lg	0
Sm	0
Grow	0
Ai	0
By	0
multi-ff	0
multi-sow	0
Feral	1

7.3.3. Surveillance components for the disease

Eight epidemiologically appropriate surveillance components were designed and simulated for their ability to achieve Austeria detection within a year of an incursion (Table 7-18). They were (i) slaughter pigs (SP): active surveillance by testing slaughter age pigs at abattoirs; (ii) sows on farms (SOF): active surveillance among breeding sows by visiting farms to test the sows; (iii) feral pig bleeding (FPB): voluntary reporting and blood sample collection on filter paper by hunters, followed by serological investigation; (iv) random lifestyle (RL): active surveillance among backyard and small

commercial pig farms; (v) risk-based 1 (RB1): risk-based surveillance option 1 targeting at domestic pig herds in the zone perceived with "the highest risk" for Austeria in Queensland, zone 8 (shown in Figure 7-1); (vi) risk-based 2 (RB2): risk-based surveillance option 2 targeting domestic pigs in the second highest risk zone present in Queensland, zone 6 (shown in Figure 7-1); (vii) farmer reporting (FR): passive disease reporting by farmers based on detection of unusual clinical signs in their domestic pigs and (viii) human case reporting (HCR): reporting of unusual cases of human disease by medical practitioners .

Table 7-18. Subjects for each Austeria surveillance component

Component code	back yard	Feral	grow	lg	multi-ff	multi-sow	Small farm
SP			�				
SOF				�რ			
FPB		▮					
RL	▮						▮
RB1	▮		▮				
RB2	▮		▮				
FR	▮		▮				
HCR	▮		▮				

Note: SP stands for slaughter pigs, SOF for sows on farms, FPB for feral pig bleeding, RL for random lifestyle, RB1 for risk-based option 1, RB2 for risk-based option 2, FR for farmer reporting, HCR for human case reporting (HCR).

Each surveillance component was implemented with three levels of sampling intensity and three levels of diagnostic sensitivity, creating 9 sub-components per component. The key assumptions on technical parameters are summarized in Table 7-19.

Table 7-19. Description of surveillance components for Austeria

Component	Description	Intensity (Test events/year)	Sensitivity (Herd level)
SP	ELISA testing of 30/20/10 pigs per herd collected randomly from healthy pigs in abattoir , 0.5/1/2 times per year	High (762) Medium (381) Low (191)	High: 80% Medium: 60% Low: 50%
SOF	ELISA testing of 30/20/10 pigs per herd collected randomly from healthy pigs on farm, 4/6/12 times per year	High (2,892) Medium (1,446) Low (964)	High: 80% Medium: 60% Low: 50%
FPB	ELISA testing of 2.5% of feral family groups each year based on harvest of 1 pig from 2.5% of the available feral families	High (1,073) Medium (537) Low (107)	High: 80% Medium: 60% Low: 40%
RL	ELISA testing of 10% of the combination of BY and SM commercial farms (total n=1850, 185 sampled), one time per year	High (370) Medium (185) Low (93)	High: 90% Medium: 70% Low: 50%
RB1	ELISA test on all commercial and backyard herds (in CRC zone 8), 4/6/12 per year	High (3,624*/2,256**) Medium (1,812/1,128) Low (1,280/752)	High: 90% Medium: 70% Low: 50%
RB2	ELISA test on all commercial and backyard herds (in CRC zone 6), 4/6/12 per year	High (288/1,236) Medium (144/618) Low (96/412)	High: 90% Medium: 70% Low: 50%
FR	Farmer recognizes clinical sign(s) and reports	High (39,694/83,877) Medium(26,463/55,918) Low (13,231/27,959)	High: 40% Medium: 20% Low: 10%
HCR	Observation, reporting, and confirmation testing based on clinical signs consistent with the disease	High: 61,785 Medium: 20,959 Low: 8,238	High: 10% Medium: 7.5% Low: 5%

Note: * number of pig herds from commercial farms; ** number of pig herds from backyard farm. SP stands for slaughter pigs, SOF for sows on farms, FPB for feral pig bleeding, RL for random lifestyle, RB1 for risk-based option 1, RB2 for risk-based option 2, FR for farmer reporting, HCR for human case reporting (HCR).

Subject enterprise types under each surveillance component are documented in Table 7-19. An example on how the intensity for a given surveillance component/sub-component was calculated is illustrated in Table 7-20.

Table 7-20. Two examples of defining surveillance intensity and probability of detection for the high intensity sub-component

Key steps	Row ID	SP	FPB	Explanation
Number herds in population	A	381	429,262	
Proportion of herds sampled per year	B	100%	1%	
Maximum test per year for a given herd	C	2	1	
Sampling compliance ratio	D	100%	25%	
Average number of testing events per year	E	762	1073	A*B*C*D
Daily probability of sampling	F	0.0055	6.8E-06	C*D/365
Herd level diagnostic sensitivity	G	90%	80%	
Daily probability of detection	H	0.005	5.5E-06	F*G

7.4. Results

7.4.1. Temporal and spatial distribution of Austeria

The temporal and spatial distribution of Austeria after it is introduced are illustrated in Figure 7-3 and Figure 7-4. A median-sized Austeria outbreak would affect 2,524 pig herds and a large size Austeria outbreak would infect more than 4,500 pig herds within a year (including feral pig families).
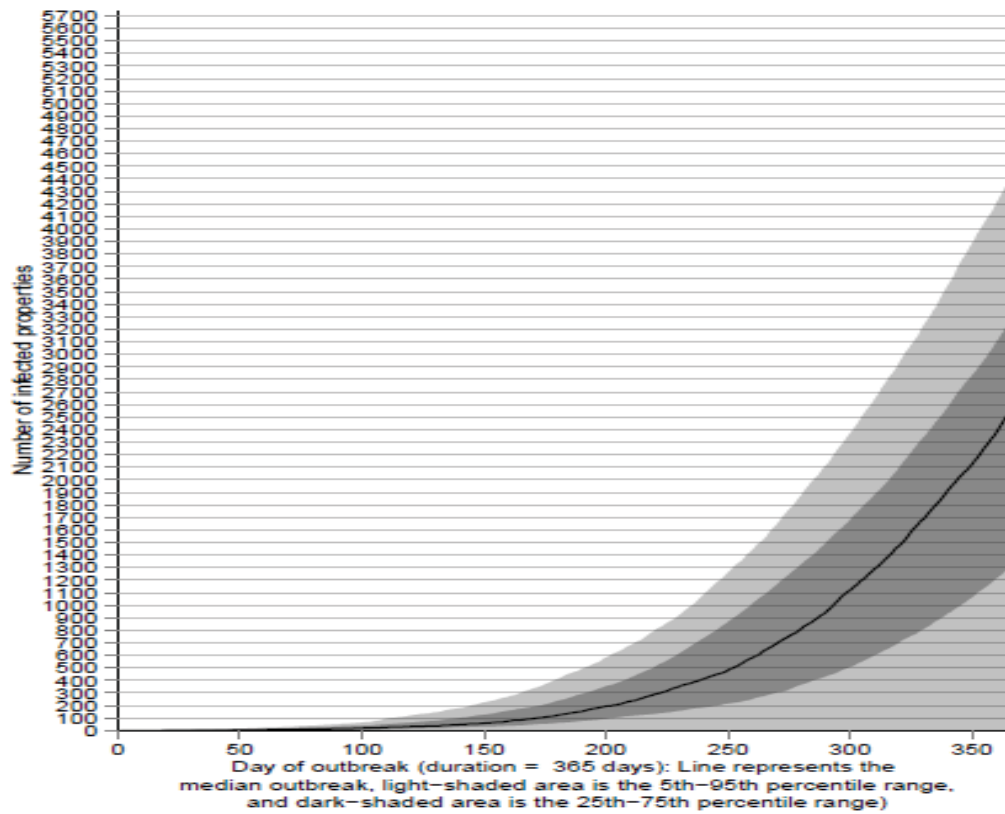

**Figure 7-3. Number of daily cumulative infected farms of an uncontrolled Austeria outbreak in Queensland, Australia**
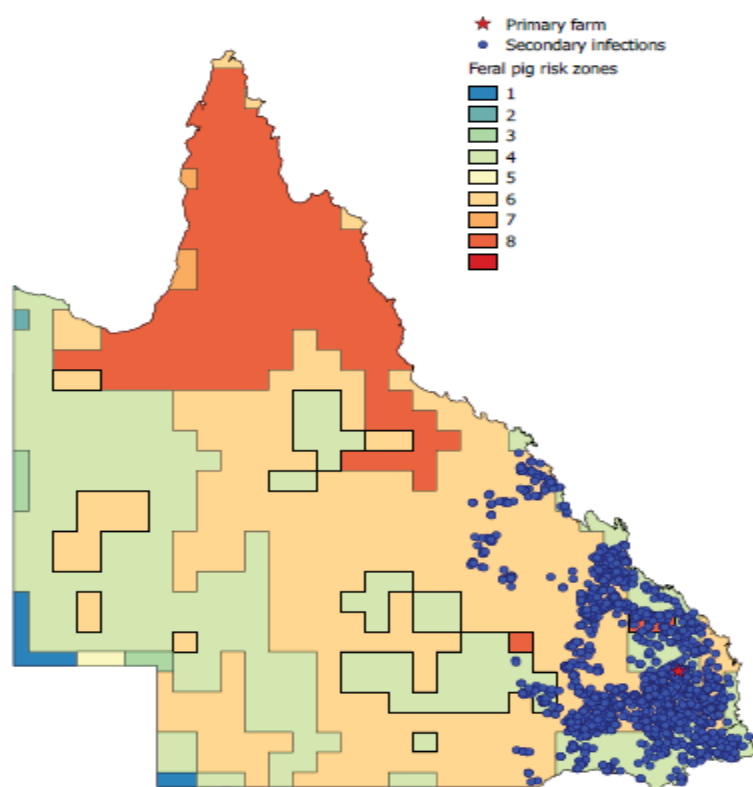
Figure 7-4. Spatial distribution of Austeria infected pig herds in a median outbreak overlaid on top of the feral pig zones in Queensland, Australia

7.4.2. Number of days taken for Austeria detection

RB2 component stands out as the most efficient surveillance option since it takes on average 137 days to detect the simulated Austeria outbreaks. The least efficient component is FPB, on average it takes 286 days to detect the simulated outbreaks (Table 7-21).

Table 7-21. Mean number of days taken for each surveillance component to detect Austeria over up to 99 iterations, Queensland, Australia (summarized for 9 sub-components)

Surveillance component	Mean	SD
RB2	137	12
RB1	197	9
RL	198	30
FR	202	33
SP	206	22
HCR	208	5
SOF	209	13
FPB	286	7

Note: SP stands for slaughter pigs, SOF for sows on farms, FPB for feral pig bleeding, RL for random lifestyle, RB1 for risk-based option 1, RB2 for risk-based option 2, FR for farmer reporting, HCR for human case reporting (HCR).

197

Figure 7-5 show results for those outbreaks that are detected in less than 365 days. The RB2 component takes the shortest median time for detecting Austeria outbreaks across all sub-components while the FPB component takes the longest. The RB2 subcomponent with high intensity and high sensitivity takes a median number of 106 days to detect the simulated outbreak, while the FPB sub-component with high intensity and high sensitivity takes a median number of 316 days to detect the simulated outbreaks.

The number of days taken for Austeria outbreak detection also varies among the sub-components within each component that differ in surveillance intensity and sensitivity. In general, the ones with high intensity and high sensitivity options tend to be the best performers that would take the shortest time to detect Austeria, while the ones with low intensity and low sensitivity take the longest time. One exception is for the FPB component. The low intensity and low sensitivity sub-component of the FPB seems to detect the disease earlier than other sub-components when measured by the median number of days taken for detection. The explanations for this result are detailed in Section 7.5.2.
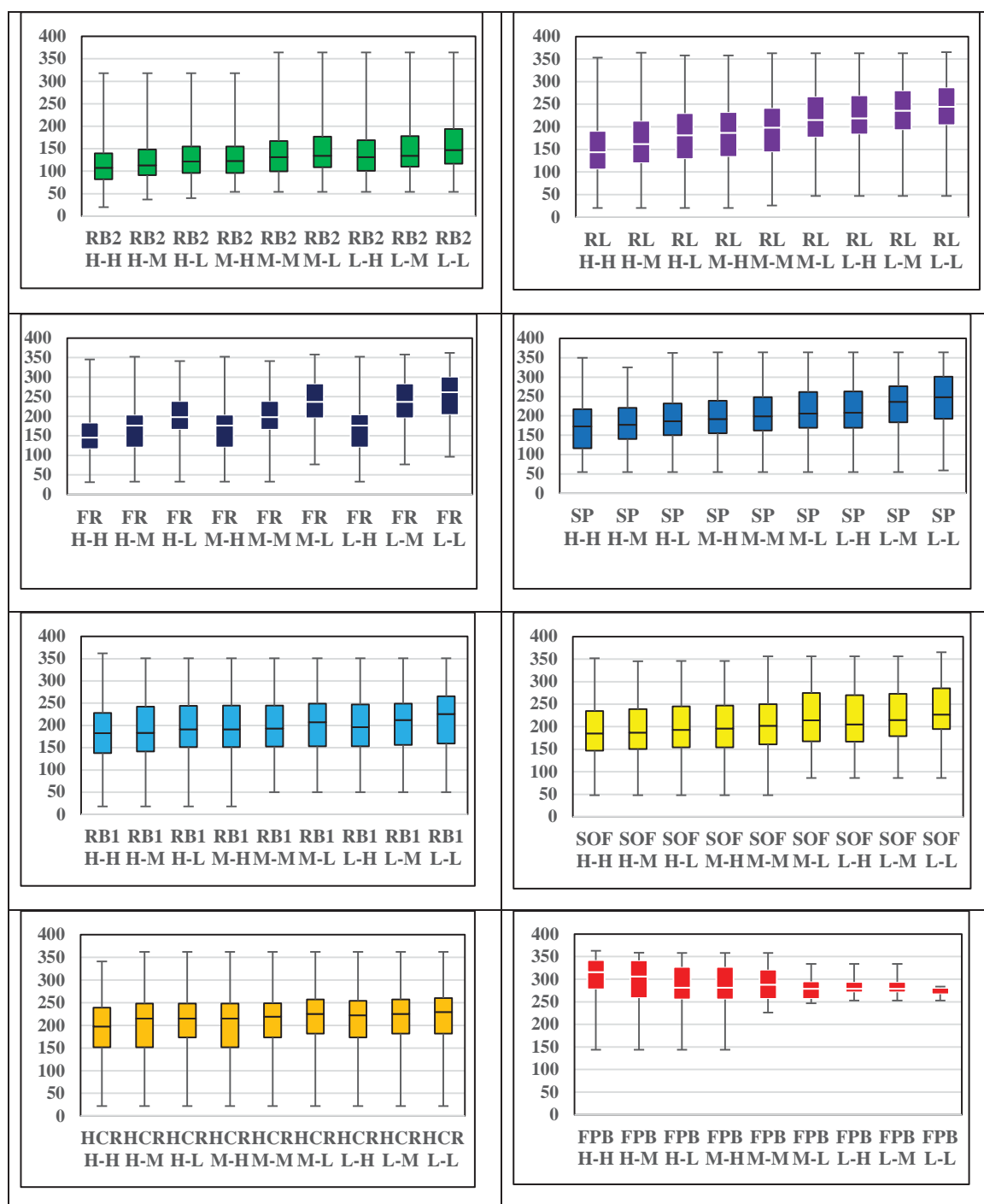
Figure 7-5. Number of days taken for Austeria detection by surveillance approach
Note: Y-axis represents number of days taken for outbreak detection. SP stands for slaughter pigs, SOF for sows on farms, FPB for feral pig bleeding, RL for random lifestyle, RB1 for risk-based option 1, RB2 for risk-based option 2, FR for farmer reporting, HCR for human case reporting (HCR). H-H means high intensity, high sensitivity; H-M, high intensity, median sensitivity; H-L, high intensity, low sensitivity; M-H, median intensity, high sensitivity, M-M, median intensity, median sensitivity; M-L, median intensity, low sensitivity; L-H, low intensity, high sensitivity; L-M, low intensity, median sensitivity and L-L, low intensity, low sensitivity.

7.4.3.Number of infected properties by the time of detection

As described before, a median-sized undetected Austeria outbreak would infect 2,524 pig herds (including feral pig families) over a period of 365 days. The most efficient surveillance component is

199

RB2, and least efficient one is FPB (Table 7-22). Component RB2 on average could detect the outbreak when only 54 owned pig herds are infected. On the contrary, Component FPB could only detect the outbreaks when 2129 owned pig herds are infected.

Table 7-22. Mean number of farms infected by the day of detection over up to 99 iterations, Queensland, Australia (summarized for 9 sub-components)

Surveillance component	Mean	SD
RB2	54	17
RB1	226	45
RL	240	168
FR	281	238
SOF	357	86
HCR	366	50
SP	390	152
FPB	2129	161

Note: SP stands for slaughter pigs, SOF for sows on farms, FPB for feral pig bleeding, RL for random lifestyle, RB1 for risk-based option 1, RB2 for risk-based option 2, FR for farmer reporting, HCR for human case reporting (HCR).

Sub-component RB2 with high intensity and sensitivity, could detect the disease when there are only 23 farms infected in the median outbreak. The least efficient one, FPB with low intensity and low sensitivity could only detect the disease when 2,515 farms are infected in the median outbreak (99.6 percent of 2,524) (Figure 7-6).
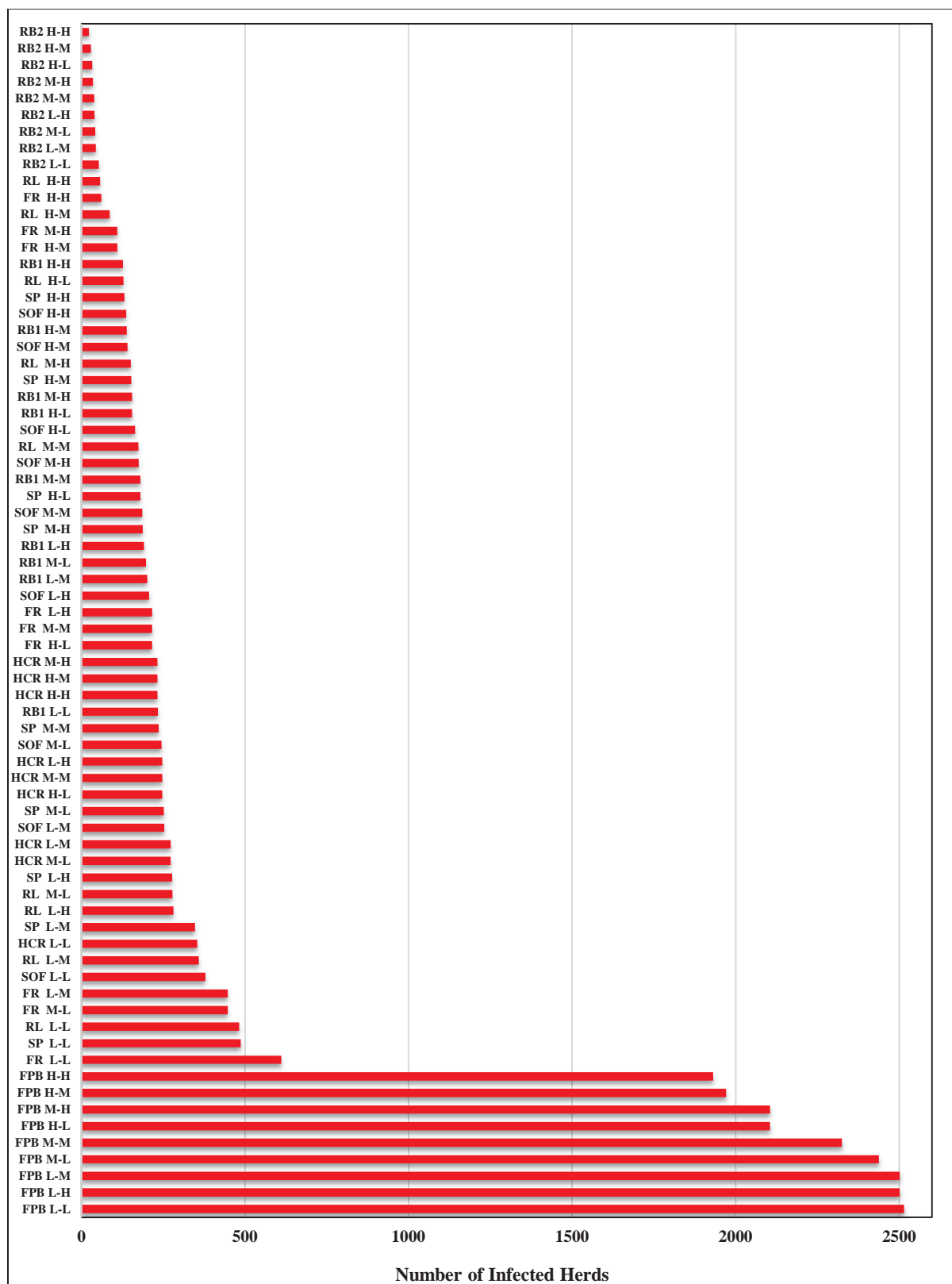
Figure 7-6. Median number of herds infected by the day of detection, by surveillance approach
Note: SP stands for slaughter pigs, SOF for sows on farms, FPB for feral pig bleeding, RL for random lifestyle, RB1 for risk-based option 1, RB2 for risk-based option 2, FR for farmer reporting, HCR for human case reporting (HCR). H-H means high intensity, high sensitivity; H-M, high intensity, median sensitivity; H-L, high intensity, low sensitivity; M-H, median intensity, high sensitivity, M-M, median intensity, median sensitivity; M-L, median intensity, low sensitivity; L-H, low intensity, high sensitivity; L-M, low intensity, median sensitivity and L-L, low intensity, low sensitivity.

201

7.4.4. Number of Austeria outbreaks detected

All RB2 subcomponents, on average, can detect 93 percent of the simulated outbreaks while all FPB sub-components detect only 15 percent of the simulated outbreaks (Table 7-22).

Table 7-23. Mean Percentage of Austeria outbreaks detected over 99 iterations for each surveillance component in Queensland, Australia (summarized for 9 sub-components)

Surveillance component	Mean	SD
RB2	93%	0.3%
RL	89%	3.4%
RB1	88%	3.8%
FR	87%	5.7%
HCR	87%	1.4%
SOF	86%	1.2%
SP	85%	2.0%
FPB	15%	11.1%

Note: SP stands for slaughter pigs, SOF for sows on farms, FPB for feral pig bleeding, RL for random lifestyle, RB1 for risk-based option 1, RB2 for risk-based option 2, FR for farmer reporting, HCR for human case reporting (HCR).

A closer look reveals that a majority of surveillance approaches are able to detect the outbreak in more than 80 percent of 99 simulated outbreaks within 365 days (Figure 7-7). The FR component with high intensity and high sensitivity is able to detect the disease outbreak in 95 percent of the simulated outbreaks while FPB with low intensity and low sensitivity can only detect 3 percent, or 3 simulated Austeria outbreaks. All RB2 sub-components can detect more than 90 percent of 99 simulated Austeria outbreaks.
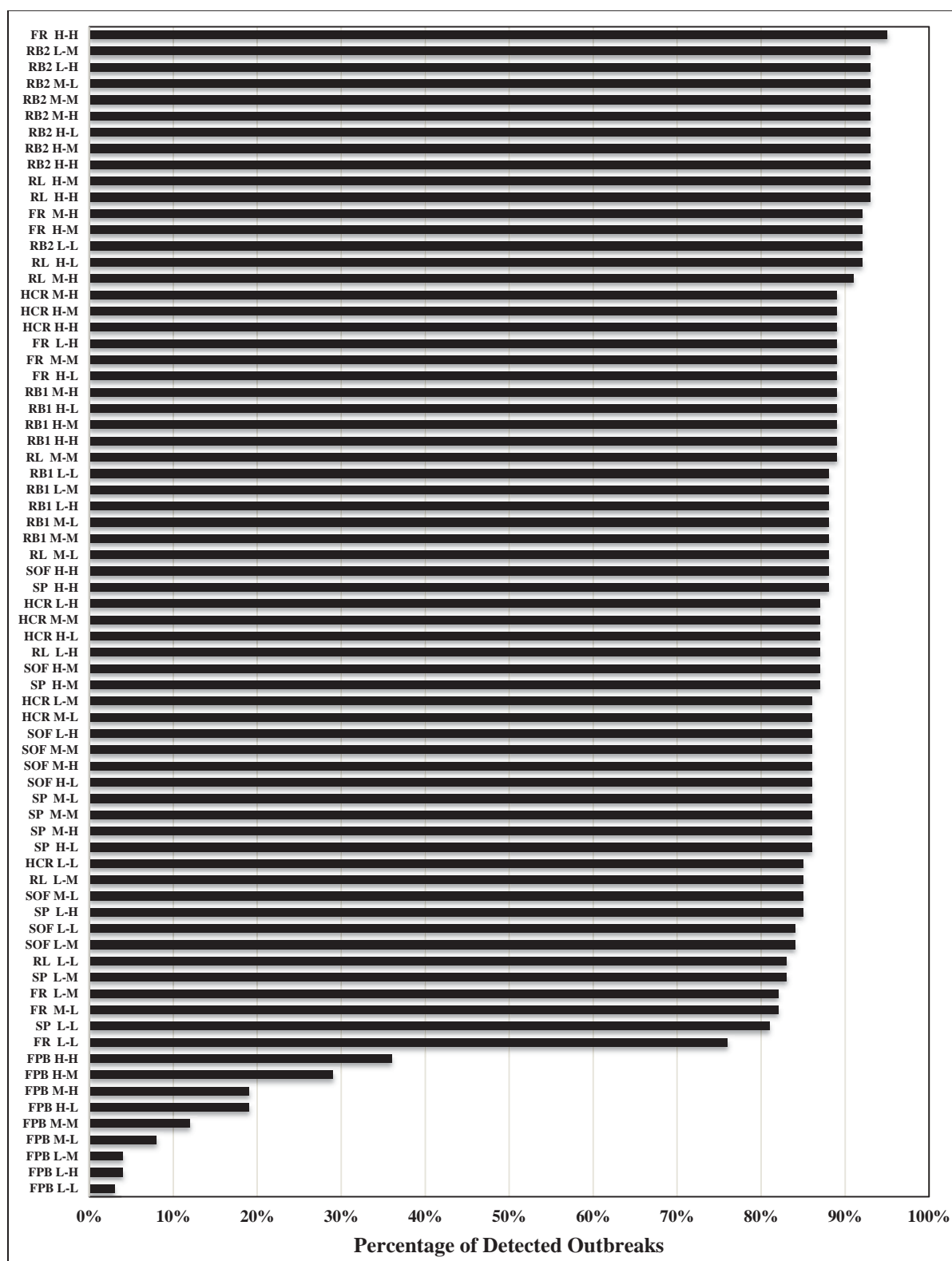
Figure 7-7. Percentage of the 99 simulated Austeria outbreaks detected by each surveillance approach within 365 days

Note: SP stands for slaughter pigs, SOF for sows on farms, FPB for feral pig bleeding, RL for random lifestyle, RB1 for risk-based option 1, RB2 for risk-based option 2, FR for farmer reporting, HCR for human case reporting (HCR). H-H means high intensity, high sensitivity; H-M means high intensity, median sensitivity; H-L, high intensity, low sensitivity; M-H, median intensity, high sensitivity, M-M, median intensity, median sensitivity; M-L, median intensity, low sensitivity; L-H, low intensity, high sensitivity; L-M, low intensity, median sensitivity and L-L, low intensity, low sensitivity.

7.5. Discussion

7.5.1. Selection of disease and location for modelling

Feral pigs can act as either a vector or a reservoir species for a number of pathogens affecting animals and human beings (Hone & Pech, 1990; Meng, Lindsay, & Sriranganathan, 2009; Naranjo, Gortazar et al., 2008; Pullar, 1950; Yob, Field et al., 2001), and can therefore contribute to disease spread following an incursion, complicating control. They were thought to have been brought to Australia by European settlers two centuries ago as domestic pigs and subsequently escaped into the wild on many occasions (Pullar, 1953). Their polyoestrous nature, together with an omnivorous diet and active foraging habit makes the species highly adaptable to a wide range of environmental and climatic conditions. Feral pigs nowadays inhabit 38 to 45 percent of Australia and are thought to be most abundant in Queensland, New South Wales, and Northern Territory (McLeod & Norris, 2004; West, 2008). Studies in Australia and elsewhere have shown that feral pigs are willing to come into close proximity to commercial piggeries, which presents the possibility of disease transmission between the two populations by short distance aerosol or direct contact if adequate barriers are not maintained (Bengsen, Gentle et al., 2014; Wu, Abril et al., 2011; Wyckoff, Henke et al., 2009).

The reason for selection of Queensland for this study was because it is one of the major pig producing states and the estimated density for both backyard pigs and feral pigs was among the highest in Australia (Figure 7-8 and Figure 7-9). In many locations within the State, the habitats of feral pigs are in the vicinity of backyard pig farms, so it would be highly possible for the pigs of the two groups to get into contact, with the greatest degree of overlapping of the populations being in Queensland.
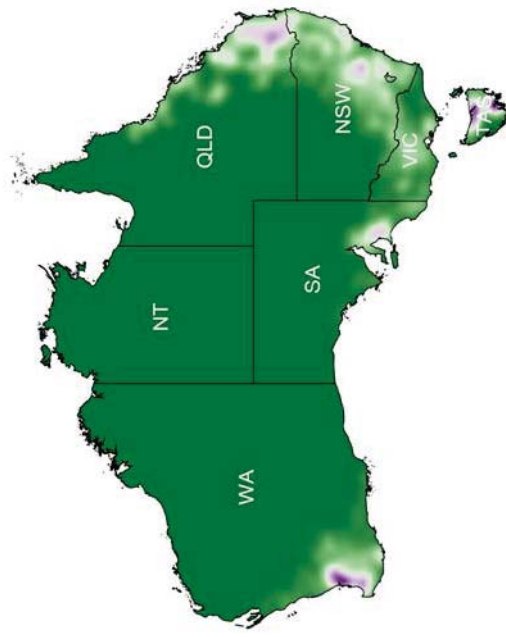
Figure 7-8. Estimated density of backyard pig herds in Australia [1]

Figure 7-9. Estimated density of feral pigs in Australia based on habitat suitability[2]

Note: 1. Estimated by Neumann et al (2012) based on the known or best estimated locations of herds identified through multiple data sources (low density, green colors, high density, purple colors); 2, data from the Invasive Animal Cooperative Research Centre, www.invasiveanimals.com ((West, 2008). Suitability has been scaled from 1 (poor suitability, green colors) to 9 (high suitability, purple colors).

205

7.5.2.Epidemiological performance of surveillance approaches

We used mean/median number of days taken for outbreak detection, mean/median number of farms infected by the day of detection and number/percent of the 99 simulated outbreaks in which disease was detected to compare the performance of the surveillance approaches tested. The results for the surveillance approaches tested are mainly consistent between the three evaluation criteria: RB2 stands out as the most successful surveillance component and FPB is the least successful. The rest of the surveillance approaches, falling between RB2 and FPB, are clustered in the middle of the ranges for the various evaluation criteria and are not substantially different in their technical performance (Table 7-21, Table 7-22 and Table 7-23).

Our modelling work demonstrated that even the most successful surveillance component, on average, would take 137 days to detect an Austeria outbreak (7-21). However, the median time to detection for the RB2 approach (131 days) is about a week shorter than the mean that is presented in the table, because the distribution of time to detection for all surveillance components is approximately lognormal. This produces a median time to detection shorter than the mean, because in a small number of outbreaks where there was a long delay to detection. The box and whisker plots in Figure 7-5 demonstrate this clearly, with the size of the "box" (25th to 75th percentile) being smaller for the more successful approaches, but the "whisker" (shortest to longest detection time) showing much less variation between approaches, both within and between components. Therefore the median time to detection is a more informative guide to selecting a surveillance approach than the mean.

The FPB sub-component with low intensity and low sensitivity seemingly outperforms the other sub-components under the same component by achieving a lower median number of days taken for detection (Figure 7-5). However, this is largely an artifact because the FPB sub-components differ in the number of simulated outbreaks detected. The FPB sub-component with high intensity and high sensitivity detected 36 percent of 99 simulated outbreaks, and the one with low intensity and low sensitivity detected only 3 simulated outbreaks of 99. Therefore because FPB is such a poor detection method with the parameter values used, the results for sub-components largely reflect differences in numbers of outbreaks detected, rather than efficiency of detection.

Surveillance component RB1 provides a good illustration of the importance of considering all epidemiological factors in deciding on surveillance methods. In theory, sampling all commercial and backyard herds in the areas with highest density of feral pigs (zone 8) seems an excellent surveillance approach, but the model results showed that under the circumstances in Queensland it is poorly effective, and sampling commercial and backyard herds in the second highest density zone (component RB2, operating in zone 6 because there are no commercial or backyard herds in zone 7 in Queensland) is a much better option, and is in fact the best component. The reason is that there is very little overlap between zone 8 (largely in the northern part of Queensland) and commercial herds, with only 24

commercial and 103 backyard herds in the zone, a total of 127 being sampled. In contrast, zone 6 extends to much more southern areas of the State and has 278 commercial and 837 backyard herds (total 1,115), so sampling these herds was a very effective method of detecting the disease. This demonstrates the benefit of modelling surveillance strategies to identify their value.

None of the surveillance strategies could detect the outbreak until it was already well established, because for at least 12 weeks the outbreak was largely limited to backyard herds and was spreading slowly before it began to increase rapidly in both backyard and feral pig herds from week 16 (Figure 7-10). The only surveillance component which achieved a median days to detection of under 16 weeks was RB2. Backyard herds are numerous but difficult to find since they are unregistered. Sampling strategy RL targeted these herds plus small commercial herds, but since there are 1,850 herds in this group across the whole of Queensland, and the strategy involved sampling only 5 to 20% of them, the H-H subcomponent (sampling 20%) detected 93% of outbreaks with a median days to detection of 144 days and a median of only 57 herds infected, but sampling 10% or 5% increased median days to detection to 161 or 182 respectively, because the probability of selecting an infected herd for sampling was considerably lower.
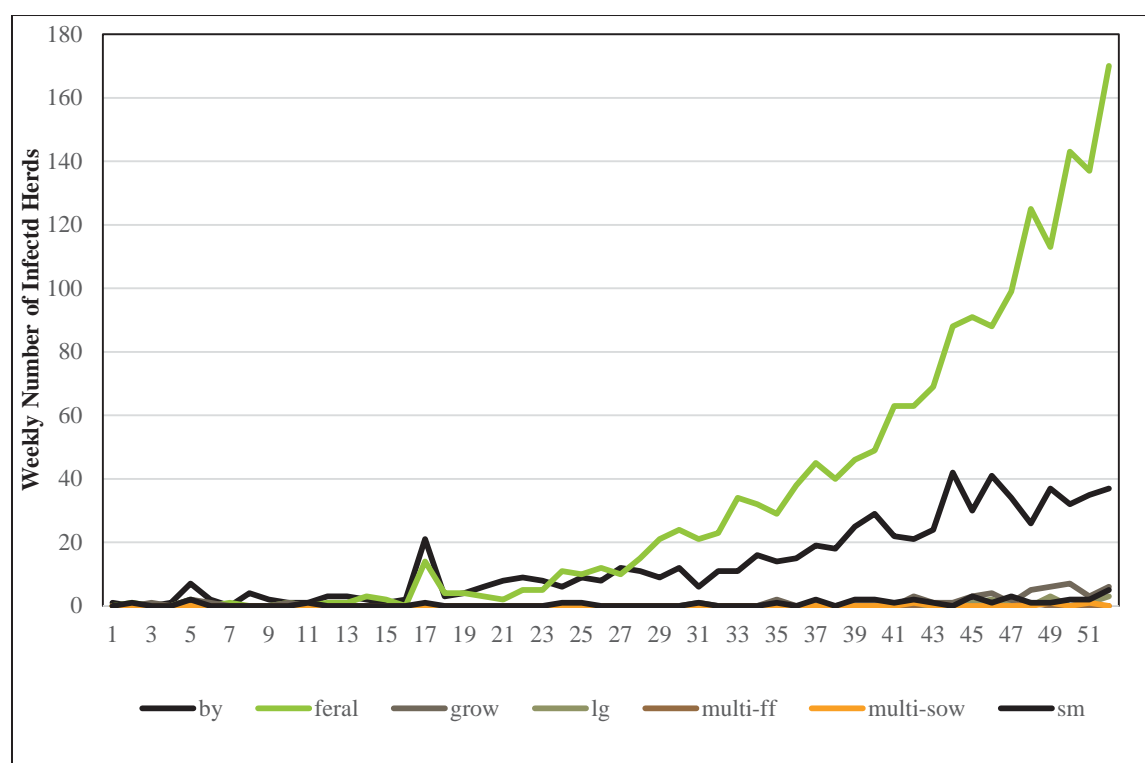
Figure 7-10. Weekly cumulative Austeria epidemic curves by pig production sector
Note: by means backyard farms unregistered (farrow to finish); grow, commercial growers and multi-site growers only (no breeders); lg, large commercial breeders (farrow to finish); multi-ff, multi-site commercial farms (farrow to finish); multi-sow, multi-site commercial breeders only; sm, small registered commercial breeder farms (farrow to finish).

Farmer reporting would be a very effective method of detection if both recognition of a novel disease and willingness to report the disease were at a high level. In this evaluation, the H-H subcomponent of FR achieved median detection at 146 days, when only 61 herds were infected, and it detected 95% of the 99 outbreaks. However, practical experience in many disease outbreaks around the world demonstrates that farmer reporting usually detects the disease quite late, due to failure of either disease recognition or willingness to report, or both – especially where infection has established in backyard herds initially. A very relevant demonstration of this has been the first of three incursions of porcine reproductive and respiratory syndrome in South Africa (Morris, 2012). In this case PRRS spread over an uncertain but extended period in backyard herds before eventually being transmitted to a commercial herd, at which point it was identified.

Reporting of human clinical cases (HCR component) is a highly efficient approach with a daily disease detection probability ranging from 0.02 to 0.08 that is among the highest in all eight surveillance components. Since we have hypothesized that Austeria was not transmitted from an infected pig to human efficiently, there would be few human cases. For a median sized Austeria outbreak, spillover into humans did not occur until day 174 with a total of 24 human infections over 365 days. HCR even overruns the RB2 component with high intensity and diagnostic sensitivity when measuring the difference between the day of disease onset and the day of detection (Table 7-24). Reporting of human clinical cases (HCR component) with high intensity and high diagnostic sensitivity could identify the disease in humans within one month of the first human case occurring, on day 208. In comparison, RB2 with high intensity and high diagnostic sensitivity subcomponent would detect the disease in pigs more than 3 months after the first farm infected, on day 108. Two policy implications of such a surveillance strategy implemented in humans need to be highlighted. Firstly, for a zoonotic disease like Austeria, reporting of human case(s) might indicate the disease, sometime on an even larger scale, are occurring among animal(s). For some diseases, reporting of sporadic human clinical case even predated any reporting on the same disease outbreaks among animals (Minh, Schauer et al., 2009). Secondly, surveillance among (sentinel) animals could provide early warning for upcoming human infections (Kulasekera et al., 2001).

Table 7-24. Comparison of HCR and RB2 by the day of onset and the day of detection

Surveillance component	Day of onset of Austeria	Day of first detection
RB2 H-H	10*	108
HCR H-H	174**	208

Note: *the first secondary infection in pigs after the index case; **the day of first human infection.

7.5.3.Contributions and limitations of our approach

The methodology presented in this paper comprises two steps: (i) development of a temporal and spatial dynamic disease model; (ii) applying several different surveillance approaches, each varying in intensity and sensitivity, to model their potential effects in terms of timeliness of disease detection, proportion of outbreaks detected as well as number of farms infected by the date of detection. Such a methodology can be generalized and used for other infectious diseases and their surveillance programs.

All the three outcome measures, in particular, the time (days) taken for detection and the proportion of outbreaks detected, are highly policy relevant. For a disease outbreak, earlier detection means a reduced loss of productivity in animals and in the case of a zoonosis, less effect on human health. For instance, late detection when using the FPB (low intensity and low sensitivity) approach, would generate, under a median Austeria outbreak, 294 more commercial farms and 6 more people infected compared to that of RB2 (high intensity and high sensitivity) (Table 7-25). Until recent years, few published studies measured or evaluated surveillance approaches by the time taken to detect a disease outbreak (Jajosky & Groseclose, 2004; Wagner, Tsui et al., 2001). The proportion of outbreaks detected *de facto* measures the sensitivity of surveillance approach under different scenarios. An effective surveillance approach, such as the RB2 subcomponents, should be able to detect almost all possible outbreaks, particularly the ones with low incidence. When assessing the technical performance of a surveillance approach by modelling, we recommend these two parameters be used jointly.

Table 7-25. Comparison of number of infected farms under different surveillance approaches

Surveillance	Days to Detection (median)	by	grow	lg	multi-ff	multi-sow	sm	Human
				Number of Infected				
RB2 H-H	108	24	3	0	0	0	2	0
FPB L-L	281	302	8	0	0	0	13	6

Note: by means backyard farms unregistered (farrow to finish); grow, commercial growers and multi-site growers only (no breeders); lg, large commercial breeders (farrow to finish); multi-ff, multi-site commercial farms (farrow to finish); multi-sow, multi-site commercial breeders only; sm, small registered commercial breeder farms (farrow to finish).

The study has identified the risk-based surveillance component RB2 as the most successful single surveillance approach by all three epidemiological criteria used. However, it fails to answer whether a combination of different surveillance approaches could produce a better result. Would RB2 surveillance approach remain the best option in case the economic aspects of each component and its alternative sub-components, and the benefit to both pig production and human health of achieving earlier detection of a disease incursion were taken into consideration? As we know, in the real world, it is seldom for a country to employ a single surveillance approach for disease surveillance, in particular

for disease outbreaks (Grein, Kamara et al., 2000; Heymann & Rodier, 2001; Morse, 2007; Wagner, Tsui et al., 2001).

Further, before exploring how to identify the most desirable overall surveillance approach, it is legitimate and critical to understand the expectations of the policy makers on the surveillance program for a given disease. Are they aiming at the one with the shortest expected time for outbreak detection or the one that could detect the highest proportion of possible disease outbreaks? Further, in addition to the epidemiological indicators of success, if cost and benefit of different options are also brought into the decision-making equation, will they together change the final selection of surveillance approaches?

7.6. Conclusion

In this article, we explore a novel approach in assessing epidemiological performance of surveillance strategies: (i) a temporal and spatial dynamic model was developed to evaluate surveillance options; (ii) a number of surveillance components and sub-components with different investigation intensity and detection sensitivity were hypothesized and tested for their performance, as measured by three epidemiological indicator variables. The findings from this work could be used to explore an epidemiologically and economically optimal surveillance portfolio, which incorporates multiple synergistic elements.

7.7. References

Barrett, R., C. W. Kuzawa, et al. (1998). "Emerging and re-emerging infectious diseases: the third epidemiologic transition." Annual review of anthropology: 247-271.

Bengis, R., F. Leighton, et al. (2004). "The role of wildlife in emerging and re-emerging zoonoses." Revue Scientifique et Technique-Office International des Epizooties **23**(2): 497-512.

Bengsen, A. J., M. N. Gentle, et al. (2014). "Impacts and management of wild pigs Sus scrofa in Australia." Mammal Review **44**(2): 135-147.

Box, G. E., G. M. Jenkins, et al. (2013). Time series analysis: forecasting and control, John Wiley & Sons.

Brahmbhatt, M. (2006). Economic impacts of avian influenza propagation. First International Conference on Avian Influenza in Humans.

Brown, C. (2004). "Emerging zoonoses and pathogens of public health significance--an overview." Revue Scientifique et Technique-Office International des Epizooties **23**(2): 435-442.

Cowled, B. D., F. Giannini, et al. (2009). "Feral pigs: predicting future distributions." Wildlife Research **36**(3): 242-251.

Daszak, P., A. A. Cunningham, et al. (2000). "Emerging infectious diseases of wildlife--threats to biodiversity and human health." Science **287**(5452): 443-449.

Fan, E. X. (2003). "SARS: economic impacts and implications."

Goutard, F. L., M. Paul, et al. (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.

Grein, T. W., K. Kamara, et al. (2000). "Rumors of disease in the global village: outbreak verification." Emerging infectious diseases **6**(2): 97.

Heymann, D. L. and G. R. Rodier (2001). "Hot spots in a wired world: WHO surveillance of emerging and re-emerging infectious diseases." The Lancet infectious diseases **1**(5): 345-353.

Hone, J. (1990). "How Many Feral Pigs in Australia." Wildlife Research **17**(6): 571-572.

Hone, J. (2012). Applied population and community ecology: the case of feral pigs in Australia, John Wiley & Sons.

Hone, J. and R. Pech (1990). "Disease surveillance in wildlife with emphasis on detecting foot and mouth disease in feral pigs." Journal of Environmental Management **31**(2): 173-184.

Jajosky, R. A. and S. L. Groseclose (2004). "Evaluation of reporting timeliness of public health surveillance systems for infectious diseases." BMC public health **4**(1): 29.

Jones, K. E., N. G. Patel, et al. (2008). "Global trends in emerging infectious diseases." Nature **451**(7181): 990-993.

Kaufmann, A. F., M. I. Meltzer, et al. (1997). "The economic impact of a bioterrorist attack: are prevention and postattack intervention programs justifiable?" Emerging Infectious Diseases **3**(2): 83.

Kulasekera, V. L., L. Kramer, et al. (2001). "West Nile virus infection in mosquitoes, birds, horses, and humans, Staten Island, New York, 2000." Emerging infectious diseases **7**(4): 722.

Longini, I. M., A. Nizam, et al. (2005). "Containing pandemic influenza at the source." Science **309**(5737): 1083-1087.

Martinez, L. (2000). "Global infectious disease surveillance." International Journal of Infectious Diseases **4**(4): 222-228.

McLeod, R. and A. Norris (2004). Counting the cost: impact of invasive animals in Australia, 2004, Cooperative Research Centre for Pest Animal Control Canberra.

Melton, G. B. and G. Hripcsak (2005). "Automated detection of adverse events using natural language processing of discharge summaries." Journal of the American Medical Informatics Association **12**(4): 448-457.

Meltzer, M. I., N. J. Cox, et al. (1999). "The economic impact of pandemic influenza in the United States: priorities for intervention." Emerging infectious diseases **5**: 659-671.

Meng, X., D. Lindsay, et al. (2009). "Wild boars as sources for infectious diseases in livestock and humans." Philosophical Transactions of the Royal Society B: Biological Sciences **364**(1530): 2697-2707.

Minh, P., B. Schauer, et al. (2009). "Association between human cases and poultry outbreaks of highly pathogenic avian influenza in Vietnam from 2003 to 2007: a nationwide study." Transboundary and emerging diseases **56**(8): 311-320.

Morris, R. (2012). "Biosecurity risks to New Zealand: Lessons from the South African experience with porcine reproductive and respiratory syndrome." New Zealand veterinary journal **60**(1): 84-85.

Morse, S. S. (2007). "Global infectious disease surveillance and health intelligence." Health Affairs **26**(4): 1069-1077.

Murray, C. J., T. Vos, et al. (2013). "Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010." The lancet **380**(9859): 2197-2223.

Naranjo, V., C. Gortazar, et al. (2008). "Evidence of the role of European wild boar as a reservoir of Mycobacterium tuberculosis complex." Veterinary microbiology **127**(1): 1-9.

Narrod, C., J. Zinsstag, et al. (2012). "A one health framework for estimating the economic costs of zoonotic diseases on society." EcoHealth **9**(2): 150-162.

Patz, J. A., P. Daszak, et al. (2004). "Unhealthy landscapes: policy recommendations on land use change and infectious disease emergence." Environmental health perspectives **112**(10): 1092.

Pullar, E. M. (1950). "The wild (feral) pigs of Australia and their role in the spread of infectious diseases." Australian Veterinary Journal **26**(5): 99-110.

Pullar, E. (1953). "The wild (feral) pigs of Australia: their origin, distribution and economic importance." Memoirs of the National Museum **18**: 7-23.

Taylor, L. H., S. M. Latham, et al. (2001). "Risk factors for human disease emergence." Philosophical Transactions of the Royal Society of London B: Biological Sciences **356**(1411): 983-989.

UNDG (2015). Socio-Economic Impact of Ebola Virus Disease in West African Countries - A call for national and regional containment, recovery and prevention. New York, USA, United Nations Development Programme Regional Bureau for Africa**:** 93.

Wagner, M. M., F.-C. Tsui, et al. (2001). "The emerging science of very early detection of disease outbreaks." Journal of Public Health Management and Practice **7**(6): 51-59.

West, P. (2008). Assessing invasive animals in Australia 2008, National Land & Water Resources Audit.

World Bank (2014). The economic impact of the 2014 Ebola epidemic : short and medium term estimates for West Africa. . W. B. Group. Washington, DC**:** 71.

Wu, N., C. Abril, et al. (2011). "Free-ranging wild boar: a disease threat to domestic pigs in Switzerland?" Journal of Wildlife Diseases **47**(4): 868-879.

Wyckoff, A. C., S. E. Henke, et al. (2009). "Feral swine contact with domestic swine: a serologic survey and assessment of potential for disease transmission." Journal of Wildlife Diseases **45**(2): 422-429.

Yob, J. M., H. Field, et al. (2001). "Nipah virus infection in bats (order Chiroptera) in peninsular Malaysia." Emerging infectious diseases **7**(3): 439.

8. Selection of optimal surveillance portfolios for detecting the incursion of an emerging zoonotic disease affecting pigs and human beings

8.1. Abstract

Background: Resources are scarce and their utilization has an opportunity cost. Decision making on disease surveillance strategies needs to be based upon sound epidemiological and economic principles.

Methods: A surveillance optimization program, OptiSurv, was developed and used for selecting optimal surveillance portfolios for detection of a hypothetical disease affecting domestic pig, feral pig and human populations after its introduction to Australia. To generate the inputs for the optimization process, both spread of the disease and the effects of 8 different surveillance approaches (referred to as 'components') and 9 sub-components for each were simulated using a spatial and temporal dynamic model. Each surveillance component/sub-component combination was modelled using 99 replicates, each replicate using different random number seeds. The economic impacts of the disease in pigs were estimated through comparing productivity differences between the scenarios with and without a surveillance program. Impacts in the human population were estimated by calculating disability-adjusted years (DALYs) lost plus costs of patient care. Different criteria were used for optimal surveillance program identification, depending on policy considerations such as the severity of human health risk versus overall economic benefit and cost of surveillance.

Results: The use of multi-component surveillance portfolios offers extra epidemiological and economic benefits over use of a single surveillance method. For eight possible surveillance components, each with nine different sub-components representing different combinations of intensity of investigation and detection sensitivity, 100 million different portfolios were possible. The optimization procedure uses alternative decision rules to rapidly identify the ten best portfolios in order, for each decision rule. The decision rules gave different weightings to human health effects, animal health effects, and costs of surveillance. In all cases, multi-component portfolios comprised all ten top options, ranking higher than all possible single-component portfolios.

Conclusion: The study introduced a practical and efficient way for screening all combinations of a set of selected surveillance components and identifying the most suitable combination of components to

make up a surveillance portfolio, using a range of decision rules which considered both epidemiological and economic factors. It demonstrated that multi-component surveillance programs could perform better than any single surveillance component, when evaluated for a range of 99 simulated variations of a single disease outbreak. The approach employed and the optimization tool used in the study are generalizable for optimization of surveillance programs for other infectious diseases.

8.2. Introduction

Optimization of surveillance program for a disease or a hazard has been defined as maximizing surveillance performance within given or expected budget constraints (Guo, Claassen et al., 2014; Prattley, Morris et al., 2007), or minimizing cost of a surveillance program given technical performance parameters (Hadorn & Stärk, 2008; Wang, Zeng et al., 2010). The technical performance means earlier detection of a disease or a hazard, improved sensitivity and/or specificity (Hutwagner, Browne et al., 2005). Technical approaches for surveillance program optimization have so far included: (i) stratification of data into different streams (Sparks, 2013); (ii) combination of different surveillance approaches (Hadorn & Stärk, 2008); (iii) using different detection algorithms for selection of a sub-set of the existing information providers without or with bearable loss of information (Polgreen, Chen et al., 2009; Scarpino, Dimitrov, & Meyers, 2012; Zhang, Jamal et al., 2011) and (iv) employing a risk-based approach (Prattley, Morris et al., 2007). Up to now, the published methods for surveillance program optimization have been focusing on achieving either the earliest detection or the highest sensitivity of detection, and the authors seldom discuss whether these strategies would be most economically efficient.

This chapter, together with Chapter 7 of the thesis, introduces a systematic approach to assessing the expected outcomes of various combinations of surveillance approaches (subsequently referred to as 'surveillance components') for an infectious disease affecting both wild and domestic animals and human beings. The goal was to identify an optimal surveillance portfolio which was expected to achieve the objective as specified in the decision rule selected for the particular disease situation. The optimization approach considered both time to detection (Chapter 7) and the economic effects of delayed detection (Chapter 8). Stochastic modelling was used for simulating both the disease outcomes and the associated economic costs and benefits of different combinations of surveillance components. By using an optimization algorithm, optimal surveillance portfolios were selected using a set of criteria appropriate to the disease situation.

8.3. Method and materials

8.3.1. Details of the disease and populations

Information concerning the population of owned pigs in Queensland (provided by Neumann, et al.[10]), the population of feral pigs, the epidemiology of the hypothesized zoonotic disease Austeria, and other relevant information about the case study can be found in Chapter 7.

The epidemiological outcomes of Austeria on pig and human populations were simulated by HandiSpread[11], an enhanced version of InterSpread Plus software. Again, the descriptions of the disease are elaborated in Chapter 7.

8.3.2. Surveillance components

Eight surveillance components were proposed for use to detect an incursion of Austeria in different populations as shown in Table 8-1. These components fall into three broad categories such as (i) laboratory surveillance among animals at either abattoirs or farms; (ii) animal or human disease reporting and (iii) surveillance among feral pigs. Each component comprised nine sub-components that were different in surveillance intensity and detection sensitivity (Table 8-1 and Table 8-2). Components and their sub-components could also be implemented in different combinations defined as surveillance portfolios. For each component selected for inclusion in a portfolio, only one sub-components could be selected since the sub-components are mutually exclusive.

Table 8-1. The nine surveillance components evaluated in the study and the populations in which they were applied

Component code	By	feral	grow	lg	multi-ff	multi-sow	sm
SP			�	▲	▲		
SOF				▲	▲	▲	
FPB		▲					
RL	▲						▲
RB1	▲		▲	▲	▲	▲	▲
RB2	▲		▲	▲	▲	▲	▲
FR	▲		▲	▲	▲	▲	▲
HCR	▲	▲	▲	▲	▲	▲	▲

Note: SP stands for slaughter pigs, SOF for sows on farms, FPB for feral pig bleeding, RL for random lifestyle, RB1 for risk-based option 1, RB2 for risk-based option 2, FR for farmer reporting, HCR for human case reporting (HCR). by means backyard farms unregistered (farrow to finish); grow, commercial growers and multi-site growers only (no breeders); lg, large commercial breeders (farrow to finish); multi-ff, multi-site commercial farms (farrow to finish); multi-sow, multi-site commercial breeders only; sm, small registered commercial breeder farms (farrow to finish).

---

[10] Neumann EJ, Hall WF, Morris RS, O'Leary B. The risk and consequences of PRRS virus introduction to Australia through importation of pork (Project 2011/1039.426). Australian Pork Ltd, Barton, ACT, Australia, 2013.

[11] InterSpread Plus; Massey University EpiCentre, Palmerston North, New Zealand. Available at http://www.interspreadplus.com, accessed August 22, 2015

Table 8-2. Description of surveillance components and sub-components for Austeria

Component	Description	Intensity (test events/year)	Sensitivity (herd level)	Cost (per event)
SP	ELISA testing of 30/20/10 samples collected randomly per herd from healthy pigs in abattoir.Testing frequency 2/1/0.5 times per year	H: 762 M: 381 L: 191	H: 80% M: 60% L: 50%	H: 460 M: 340 L: 220
SOF	ELISA testing of 30/20/10 samples per herd collected randomly from healthy pigs on farm. Testing frequency 12/6/4/times per year	H: 2892 M: 1446 L: 964	H: 80% M: 60% L: 50%	H: 460 M: 340 L: 220
FPB	ELISA testing of 2.5% of feral family groups each year based on harvest of 1 pig from 2.5% of the available feral families	H: 1073 M: 537 L: 107	H: 80% M: 60% L: 40%	H: 260 M: 236 L: 212
RL	ELISA testing of 10% of the combined population of BY and SM commercial farms (total n=1850, 185 sampled), once per year	H: 370 M: 185 L: 93	H: 90% M: 70% L: 50%	H: 184 M: 160 L: 136
RB1	ELISA testing of all commercial and backyard herds (in CRC zone 8) in a single random week, 12/6/4times per year	H: 3624*/2256** M: 1812/1128 L: 1280/752	H: 90% M: 70% L: 50%	H: 460/184 M: 340/160 L: 220/136
RB2	ELISA testing of all commercial and backyard herds (in CRC zone 6), 12/6/4times per year	H: 288/1236 M: 144/618 L: 96/412	H: 90% M: 70% L: 50%	H: 460/184 M: 340/160 L: 220/136
FR	Farmer recognizes some unusual clinical signs in pigs and reports them to the veterinary service	H: 39694/83877 M: 26463/55918 L: 13231/27959	H: 40% M: 20% L: 10%	H: 5.5/1 M: 2.75/0.5 L: 1.38/0.25
HCR	An affected person is recognized as having a novel disease by a medical practitioner, and reports the findings to the health services.	H: 61785 M: 20959 L: 8238	H: 10% M: 7.5% L: 5%	H: 300 M: 150 L: 115

Notes: SP stands for slaughter pigs, SOF for sows on farms, FPB for feral pig bleeding, RL for random lifestyle, RB1 for risk-based option 1, RB2 for risk-based option 2, FR for farmer reporting, HCR for human case reporting (HCR). * the number of commercial pig herds under surveillance; ** the number of backyard pig herds under surveillance. Details of the surveillance strategies are given in Chapter 7.

8.3.3.Simulation modelling of Austeria and surveillance components/sub-components

Each of the 72 surveillance component/sub-component combinations was simulated for 99 iterations of the disease, using HandiSpread, making a total of 7,128 model runs. By using 99 iterations of the model with each based on a separate random number seeding process, variability in the disease process was represented while keeping all parameter settings constant. Each simulation stopped when the disease was detected by that surveillance method or the model had run for 365 days, whichever came first. The

key outputs of each model run included: days taken for detection, number of farms infected by the day of detection, and the proportion of simulated outbreaks detected within 365 days, which were summarized across the set of 99 simulations for each component/sub-component option. These pieces of information were used to determine the economic impact of an Austeria outbreak and the benefit of detection, and this data was used when searching for the optimal surveillance portfolio(s).

8.3.4. Estimation of economic effect of an Austeria outbreak and benefit of surveillance

Definitions of the key economic metrics used for the estimation of economic impacts of the disease and benefit of a given surveillance component/sub-component are documented in Table 8-3.

Table 8-3. Terminology and definitions used in the economic evaluation of surveillance components/sub-components

Terminology	Definition
Total economic impact of Austeria	The summation of the impact on the pig industry and the impact of human cases of the disease. The impact on the pig industry is the difference between average net revenue per pig under Austeria outbreak scenario versus Austeria-free scenario multiplied by the number of pigs produced by that particular pig enterprise
Average net revenue per pig	The current market value of a pig minus the average feed cost, non-feed variable cost, fixed overhead cost of production.
Gross benefit of surveillance	Averted impact of Austeria from the date of detection for a given surveillance approach
Net benefit of surveillance	Gross benefit of surveillance minus the cost for surveillance
Benefit cost ratio	Gross benefit divided by surveillance cost by the day of detection

8.3.4.1.   Cost for surveillance

Only the variable cost for implementing any given surveillance strategy was estimated. This included the cost incurred for laboratory investigation, travel, sample collection and transportation, etc over a period of up to 365 days. We assumed the cost for laboratory investigation was 12 Australian dollars per sample for screening test. Cost for collection of samples, travel, transportation was 100 Australia dollars for SP, SOF, RL, RB1 and RB2, 200 dollars for FPB and 0 for FR and HCR. The cost per testing event was multiplied by the number of testing events. Cost for laboratory investigation of human infection is 125 Australian dollars. Zero dollar cost for human sample collection, transportation, etc. Fixed cost of maintaining the infrastructure to conduct surveillance was not taken into consideration.

8.3.4.2. Economic consequence of Austeria on the pig industry

The economic impact of an Austeria outbreak comprised two parts: (i) economic loss to pig industry and (ii) economic value of lost human life and productivity. For each surveillance component/sub-component that was evaluated, the net economic loss associated with the disease was accrued to the day of outbreak detection or 365 days, whichever came first.

A budgeting model was developed by Neumann *et* al [12] to estimate the typical cost of production and income from pig production on Australian pig farms. The results of this analysis were used in this study to estimate economic inputs and outputs from production under an Austeria-free scenario, and then this was adjusted to estimate the economic consequences of an Austeria outbreak at individual enterprise level. Expenses were partitioned across the following four stages of production: breeding and gestation, lactation (up to four weeks after parturition), early post-weaning ('weaner' stage; four to 12 weeks of age), and growing (12 to 20 weeks of age). Data on various cost elements were obtained from literature and experts' opinions (Table 8-4).

Table 8-4. Diet specifications and ingredient costs for Australian pig budgeting model

| Ingredient | $/tonne[a,b] | Diet specification | | | | | |
		Creep %	Weaner %	Grower %	Finisher %	Dry %	Lact %
Barley	265	0.00	0.00	40.50	88.56	79.60	33.00
Maize	290	0.00	0.00	0.00	0.00	0.00	0.00
Meat / bone meal	720	5.00	8.00	10.00	8.00	8.00	10.00
Dried blood meal	880	0.00	2.50	4.00	3.00	0.00	1.50
Fishmeal	2,000	7.50	4.00	0.00	0.00	0.00	2.00
Skim milk powder	1,100	25.00	5.00	0.00	0.00	0.00	0.00
Wheat	275	54.17	72.21	40.00	0.00	12.00	50.00
Synthetic lysine	3,200	0.40	0.42	0.20	0.17	0.15	0.25
Tallow	990	0.00	0.00	0.00	0.00	0.00	3.00
Soybean oil	1,370	2.50	2.50	0.00	0.00	0.00	0.00
Soybean meal	550	5.00	5.00	5.00	0.00	0.00	0.00
Salt	460	0.00	0.00	0.00	0.00	0.00	0.00
Synthetic methionine	6,400	0.18	0.12	0.05	0.02	0.00	0.00
Synthetic threonine	5,120	0.00	0.00	0.00	0.00	0.00	0.00
Premix (weaner)	12,800	0.25	0.25	0.00	0.00	0.00	0.00
Premix (grower)	12,800	0.00	0.00	0.25	0.25	0.00	0.00
Premix (sows)	12,800	0.00	0.00	0.00	0.00	0.25	0.25
Total (%)		100	100	100	100	100	100
Freight	32						
Total cost (per ton)		760	560	425	389	370	451

Note: [a]Anonymous. Eyes and Ears. Australian Pork Limited, Kingston, ACT, Australia, 2013; [b] communication with Barugh IW, Massey University, February 15, 2013; all costs are in Australia dollars.

---

[12] Neumann EJ, Hall WF, Morris RS, O'Leary B. The risk and consequences of PRRS virus introduction to Australia through importation of pork (Project 2011/1039.426). Australian Pork Ltd, Barton, ACT, Australia, 2013.

Cost of production in the base model was subtracted from the current market value of a slaughter age pig (baconer) carcass to determine the net revenue per baconer sold. Using this process, the average net revenue per pig sold was calculated across all baconers sold from a "typical" pig enterprise within each production system. For the surveillance study, a single typical value was used for each type of production system, with no consideration of variation with enterprise types. The total profit margin was then estimated by using the average net revenue per pig multiplied by the number of pigs marketed. Loss of pigs because of mortality both with and without the effect of Austeria was monetized and averaged across the pigs surviving by the end of each production stage.

We hypothesized that the effects of Austeria on pig productivity would be similar to those of porcine reproductive and respiratory syndrome (PRRS), the disease on which Austeria was modelled. The data on the effects of PRRS was based on a large recent study on the cost of the disease on the United States pig industry over a 12-month period following an outbreak of the disease in a previously negative herd (Holtkamp, Kliebenstein et al., 2007). The key assumptions for economic loss estimation are documented in Table 8-5.

Table 8-5. Key performance metrics influencing cost of production and net revenue per pig in herds with and without Austeria

	Base model			Austeria model		
	Min	Likely	Max	Min	Likely	Max
Breeding herd						
Born alive per litter	10.5	10.8	11.2	9.6	10.0	10.3
Pre-weaning mortality	9.39%	11.99%	13.39%	15.00%	15.52%	19.00%
Litters/mated female/year	2.20	2.25	2.33	2.09	2.19	2.22
Culling rate	50.24%	54.67%	59.24%	47.00%	51.49%	56.00%
Mortality rate	8.86%	10.08%	11.56%	10.20%	12.29%	12.90%
Growing herd (post-weaning)						
Mortality rate	3.60%	4.78%	6.60%	6.40%	8.01%	9.40%
Average daily gain (g/day)	667	701	727	650	680	710
Feed conversion rate	2.41	2.50	2.54	2.37	2.41	2.50
Percent sold as prime	95.4%	96.0%	97.4%	94.8%	95.8%	96.8%

The Neumann et al budgeting model was used for estimating cost of adjusted pig productivity under Austeria. The resulting change in average net revenue per pig in the 'Austeria' model multiplied by the number of pigs produced (or sold) resulted in the total cost of the outbreak for a herd.

Other important assumptions included:

- The effect per pig of Austeria infection on large breeders, small breeders, and multi-site farrow to finish breeders would be the same, as the primary farm revenue from these farm classes was generated from the sale of baconers.

- The cost of Austeria infection for multi-site farrow to feeder breeding farms was restricted only to

its effect on production up to the 12 week age at sale.

- Grower and multi-site grower farms were considered as the downstream 'customers' of multi-site breeding farms, receiving 12 week old pigs for feeding to market weight at around 20 weeks of age. Economic costs on these farm classes included expenses due to increased cost of production and loss of revenue from selling pigs at a lighter bodyweight. While the cost of PRRS on other farms was annualized based on the estimated biological productivity of the farm for a 12 month period after the infection, the cost of PRRS to an individual growing pig was annualized by multiplying the point in time inventory for the site by six to account for the fact that the population of pigs would be turned over approximately every 8 to 9 weeks, or about six times per year.

- The cost of Austeria in backyard pig (BYP) herds was calculated in a manner similar to that of very small farrow to finish herds. However, an adjustment was made to account for the fact that not only was the sale of market weight pigs an important source of revenue for a BYP herd but that production of pork meat for consumption by the owner's family was also an important farm output, and if this output was lost then equivalent quantities of meat would have to be purchased. As BYP farms relied almost exclusively on free or very low cost sources of feed such as grazing and waste food as their primary source of nutrition and labour is unpaid, the cost of production in the post-weaning phase of BYP production was considered negligible. A simple model for assigning the cost of an Austeria outbreak to a BYP was therefore established that represented the sum of 80% of the cost of Austeria on a similar-sized very small commercial farrow to finish breeding herd plus the expense associated with having to purchase 50% of their annual pork consumption.

- Feral pig family outbreaks were assigned a cost of $0.01 per family so that non-zero values could be included in calculations.

- The disease would fully run its course in 16 weeks. We assumed that after 16 weeks, the farm would return to normal productivity and the farms would have an adequate degree of herd immunity to prevent further virus circulation.

The economic effect of an individual herd outbreak of Austeria was calculated by subtracting the profit margin under the non-Austeria scenario from the profit margin under the Austeria scenario, producing a net loss per pig, which when multiplied by the number of pigs produced a herd level economic effect. The unit economic loss per pig due to Austeria in each production system is summarised in Table 8-6.

Table 8-6. Economic net loss per pig due to Austeria under different production systems in Australia

Production type	Economic loss per pig produced (A$)*
Weaned pig in weaner producing herd	11.34
Farrow to finish herd	37.5
Grower herd only	26.16
Backyard pig (cost per farm)	401=37.5*0.8 + 0.5*742

Note: * unit cost in Australia dollars. Cost for commercial pig herds is estimated per pig produced. The cost for backyard pig herds is estimated per farm.

8.3.4.3. Economic cost of Austeria on human population

The cost for human infection of Austeria comprised two parts: (i) direct cost: for treatment of affected people; (ii) indirect cost: productivity loss because of the disease.

Indirect cost of Austeria on human population was estimated in two steps. Firstly, the disease burden caused by Austeria was summarized in DALYs (Murray & Acharya, 1997) by using WHO's DALY template[13]. Secondly, the number of DALYs was converted into dollar value by using the human capital approach[14].

Key assumptions for estimation of the number of DALYs per human case included:

- The high risk group for Austeria consists of all people living in the immediate proximity of the pig farms and coming in direct contact with pigs. We assumed this was a population of 10,000 people involved in the management of all commercial and backyard pig farms, which represented an average of 4.5 persons per farm. They were considered to range in ages at the time of exposure to Austeria between 15 and 69 years old; to further simplify the calculation, we assumed only male residents got infected.

- Epidemiological and clinical characteristics of human infections: the disease was transmissible from pigs to people directly handling them, but did not cause human to human transmission. Among the people who became infected, 25 percent had mild clinical manifestations and they would suffer 50 percent disability for 1 month. Another 25 percent experienced severe symptoms and suffered 100 percent disability for a year. The remaining 50 percent died.

- Treatment cost: 250 AUD per mild case, 12,500 AUD per severe case who recovered and 10, 000 AUD per deceased case.

- The effect of loss of income was estimated using nominal GDP per capita, which was 62,000 USD

---

[13] http://www.who.int/healthinfo/global_burden_disease/tools_national/en/
[14] Health of Nations: The Value of a Statistical Life. Australian Safety and Compensation Council, Australian Government, July 2008 pp. 194.

in 2014, obtained from the World Bank Group website[15]. At that time the US and Australian dollars were at approximate parity, so this figure was also treated as the amount in Australian dollars.

8.3.4.4. Summary of the economic analysis outputs

An Excel spreadsheet was created to document the potential economic effects of each total multi-unit outbreak of Austeria, using as input the results from HandiSpread simulations on the effect of surveillance components/sub-components plus the economic calculations described above. This file was used as an input for the optimization process.

8.3.5. Surveillance optimization approach

An optimization program, called OptiSurv, was developed in Excel format, using Visual Basic programming. The process for optimization is illustrated in Figure 8-1. The program identified the optimum portfolio from a number of surveillance components (in OptiSurv, a component is called a strategy), each of which could have a number of subcomponents (called sub-strategies in OptiSurv) with varied surveillance intensity and detection sensitivity (refer to Section 8.3.2), plus an additional sub-component that corresponded to "do not include this component in the portfolio".
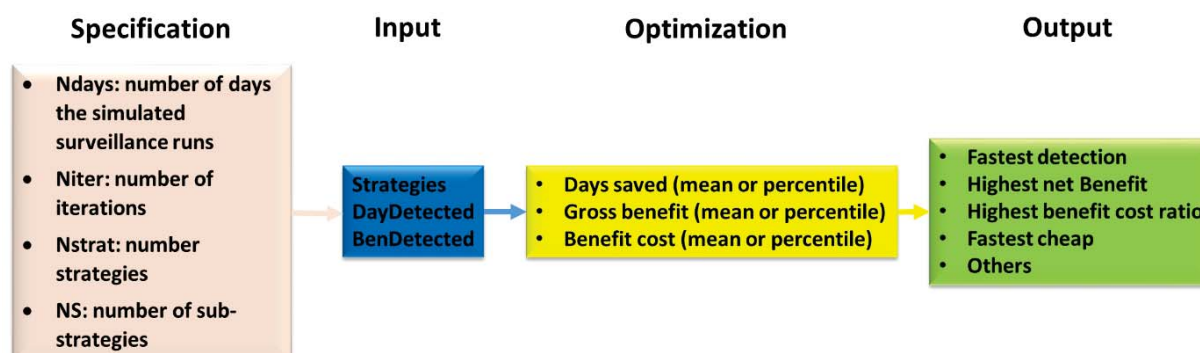
Figure 8-1. Optimization process under OptiSurv Program

OptiSurv allows you to choose from a number of alternative criteria when searching for the best surveillance portfolios as follows:

- "Fastest detection" selects the portfolio which will on average find the disease first, without consideration of economic aspects.
- "Fastest cheap" picks out low cost portfolios which still give quite rapid detection, although not as good as the first option.

- "Highest net benefit" concentrates on economic efficiency, although it can be combined with a requirement for rapid detection.
- "Highest benefit cost ratio" ensures maximum economic payoff per dollar invested, but is likely to pick very low cost strategies which may be inefficient in detecting the disease.

With eight surveillance components and nine sub-components per component, plus "no action", there were 100 million possible surveillance portfolios to be compared[16]. The optimization process within OptiSurv therefore evaluates all possible portfolios in a computationally-efficient manner that reduces processing time. It does this by systematically building up each portfolio one component and sub-component at a time, using a goal-seeking strategy which rapidly discards suboptimal portfolios, and concentrates on identifying the top 100 portfolios, which can then be ranked using the analyst's preferred decision rules based on speed of disease detection plus economic costs and benefits, as demonstrated later in this chapter.

The first step is to consider the minimum time to detection achieved for each of the 99 outbreak iterations by portfolios, progressively building each portfolio by adding component/sub-component combinations until it becomes clear whether or not it is a contender for the "short list", and if not discarding it. Each pair of component/sub-component is evaluated only once, and that result is used each time that combination is considered with other component/sub-components. It is also necessary to exclude portfolios which achieve rapid detection, but contain component/sub-components which do not contribute to detection speed. The evaluation still takes up to a few hours because of the size of the task, but it is much faster than if all possible portfolios had to be fully evaluated. With such a large number of possible portfolios, it is inevitable that differences between many similarly constructed portfolios will be slight, but some clusters of portfolios will stand out as superior to the bulk of alternatives, and they will be included in the top-ranked group. By default OptiSurv keeps track of the 100 "best" portfolios found while it is evaluating all possible portfolios in case there are alternative portfolios that are close to the optimal portfolio that might be more practical to use. Figure 8-2 shows the layout of the summary screen of OptiSurv, through which the various steps of the optimization process are managed.

---

[16] There are eight surveillance components, nine sub-components plus one"don't choice" option. The possible combinination of different options could be 10^8.

Set up a new analysis

Days:	365	Enter details about the simulated outbreaks
Strat:	8	
Substr:	72	Check the Strategies sheet
Iter'n:	99	

Create the Benefits worksheet

Import a Benefits file

Read farm cost file

Read simulated outbreaks

Delete FarmCosts and Epidemics worksheets

Clear the DayDetected worksheet

Confirm that you are happy with the number of strategies and substrategies

Fill the DayDetected worksheet

Import a table file

Read the simulated surveillance data

Delete Benefits worksheet

Find best portfolio

Costs updated	Set budget limit
Set detection	Set criterion

Find best portfolio

History of processing

Strategies	abatt_gf	monthly_breed	hunter
Levels	9	9	9

1/09/2015 11:53	C:\Users\Rob\Documents\ROBcancount\Austeria
1/09/2015 11:55	C:\Users\Rob\Documents\ROBcancount\Austeria
1/09/2015 11:56	C:\Users\Rob\Documents\ROBcancount\Austeria

Within the budget of	1000000		on average
Portfolio must improved detection by		0	days
Optimisation criterion		average	days saved
2/09/2015 7:35:08			
2/09/2015 7:50:53			

Figure 8-2 Summary screen of OptiSurv

8.4. Results

Each of the alternative decision rules will now be considered, and results presented for the top-ranked portfolios based on each rule. Factors influencing the choice of which decision rule to apply in a particular disease situation are also outlined.

8.4.1. Portfolios with the shortest expected 'number of days to detection'

This option chooses a portfolio purely on how fast the disease would be detected, regardless of cost. It may be appropriate in a case where the possibility of the disease entering the country is causing severe panic in the population, and the government instructs that no expense is to be spared in undertaking surveillance. By using such a portfolio, Austeria could be detected by day 96 after incursion, on average. Compared with no surveillance (365 days of disease spread), this represents a saving of 269

days. A number of portfolios can equally achieve the fastest detection time of 96 days. The top ten portfolios are listed in Table 8-7.

Since there is no economic element taken into consideration in the optimization process, all components except FPB are used with high or medium intensity and sensitive sub-components. Notably, none of these portfolios included an FPB surveillance component because it does not reduce time to detection any more than the other seven components together. Average gross benefit of implementing any of these ten surveillance portfolios was the same at $5,980,587 because the number of farms infected and the associated impacts were the same in all cases by the same day of detection. However these surveillance portfolios vary in benefit cost ratio and net benefit because the cost for implementing them varies from $0.95 m to over $1.1 m. The net benefit is $760,000 to 950,000 less than that of portfolios chosen using economic criteria.

8.4.2. Portfolios with the highest net benefit

Portfolios chosen solely on the basis of highest net benefit usually consisted of RL, FR and HCR surveillance components. This is because they use only low cost surveillance components. The cost for surveillance is about 6% of the cost of the portfolios that achieve fastest detection, the average time to detection was 38 days longer, allowing greater spread of disease. The best portfolio in this group could achieve the an extra of $947,337 than that of the fastest detection portfolio and the surveillance cost was about $1,063,930 lower. The top ten surveillance portfolios are presented in Table 8-8.

Table 8-7. Top ten surveillance portfolios for Austeria in Australia with the shortest days to detection

Rank	SP	SOF	FPB	RL	RB1	RB2	FR	HCR	Ave B/C	Ave NB	Ave SC
1st	H-H	H-H	NU	H-H	H-H	H-H	H-H	H-H	8.5	4847797	1132789.9
2nd	H-H	H-M	NU	H-H	H-H	H-H	H-H	H-H	9.3	4939035	1041552.0
3rd	H-H	H-L	NU	H-H	H-H	H-H	H-H	H-H	10.2	5030273	950314.2
4th	H-H	M-H	NU	H-H	H-H	H-H	H-H	H-H	10.1	5022670	957917.3
5th	H-H	H-H	NU	H-H	H-H	H-H	H-H	H-M	8.6	4849161	1131425.5
6th	H-H	H-M	NU	H-H	H-H	H-H	H-H	H-M	9.3	4940409	1040178.1
7th	H-H	H-L	NU	H-H	H-H	H-H	H-H	H-M	10.2	5031657	948930.6
8th	H-H	M-H	NU	H-H	H-H	H-H	H-H	H-M	10.1	5024053	956534.6
9th	H-H	H-H	NU	H-H	H-H	H-H	H-H	M-H	8.6	4849162	1131425.5
10th	H-H	H-M	NU	H-H	H-H	H-H	H-H	M-H	9.3	4940409	1040178.1

Note: SP stands for slaughter pigs, SOF for sows on farms, FPB for feral pig bleeding, RL for random lifestyle, RB1 for risk-based option 1, RB2 for risk-based option 2, FR for farmer reporting, HCR for human case reporting (HCR). H means either high intensity or high sensitivity, M either medium intensity or medium sensitivity and L either low intensity or low sensitivity; Ave B/C means average benefit cost ratio; Ave NB means average net benefit, in Australian dollars; Ave SC means average surveillance cost, in Australian dollars; NU means not used.

Table 8-8. Top ten surveillance portfolios for Austeria in Australia with the highest net benefit

Rank	SP	SOF	FPB	RL	RB1	RB2	FR	HCR	Ave NB	Ave DS	Ave B/C	Ave SC
1st	NU	NU	NU	H-H	NU	NU	M-H	H-L	5795134.1	231	132.9	68859.2
2nd	NU	NU	NU	H-H	NU	NU	M-H	L-H	5795134.1	231	132.9	68859.2
3rd	NU	NU	NU	H-H	NU	NU	M-H	H-M	5795015.3	232	129.6	70563.7
4th	NU	NU	NU	H-H	NU	NU	M-H	M-H	5795015.3	232	129.6	70563.7
5th	NU	NU	NU	H-H	NU	NU	M-H	M-M	5794618.1	231	132.0	69375.2
6th	NU	NU	NU	H-H	NU	NU	M-H	L-L	5794437.4	230	131.3	67426.8
7th	NU	NU	NU	H-H	NU	NU	M-H	M-L	5794408.3	231	131.3	68165.8
8th	NU	NU	NU	H-H	NU	NU	M-H	L-M	5794408.3	231	131.3	68165.8
9th	NU	NU	NU	H-H	NU	NU	M-H	H-H	5794288.4	232	126.0	72562.0
10th	NU	NU	NU	H-M	NU	NU	M-H	H-M	5789510.9	231	132.8	67925.8

Note: SP stands for slaughter pigs, SOF for sows on farms, FPB for feral pig bleeding, RL for random lifestyle, FPB for feral pig bleeding, RL for random lifestyle, RB1 for risk-based option 1, RB2 for risk-based option 2, FR for farmer reporting, HCR for human case reporting (HCR). H means either high intensity or high sensitivity, M either medium intensity or medium sensitivity and L either low intensity or low sensitivity; NU means not used; Ave NB means average net benefit; Ave DS means average days saved; Ave B/C means average benefit cost ratio; Ave SC means average surveillance cost.

8.4.3.Portfolios with the highest benefit cost ratio

It would seem from first principles that the portfolio with the highest benefit-cost ratio would be a very beneficial one. However in reality this approach chooses a minimum surveillance portfolio with a cost of $10,000 or less, which yields extremely high benefit cost ratios but delays detection by about 140 days. Most of the surveillance portfolios with high benefit cost ratio have only one surveillance component, HCR, which has minimal cost, while a few portfolios include RL and FR as well (Table 8-9). These portfolios rely on detection and reporting of human cases, and are only likely to be rational for a minor disease which has a low impact on human health.

Table 8-9. Top ten surveillance portfolios for Austeria in Australia with high benefit cost ratio

Rank	SP	SOF	FPB	RL	RB1	RB2	FR	HCR	Ave B/C	Ave DS	Ave GB	Ave NB	Ave SC
1st	NU	NU	NU	NU	NU	NU	NU	L-L	1859.2	124	4498575.9	4494843.6	3732.3
2nd	NU	NU	NU	NU	NU	NU	NU	M-L	1388.6	131	4691094.4	4685672.1	5422.3
3rd	NU	NU	NU	NU	NU	NU	NU	L-M	1388.6	131	4691094.4	4685672.1	5422.3
4th	NU	NU	NU	NU	NU	NU	NU	H-L	1168.0	138	4865743.4	4858718.9	7024.6
5th	NU	NU	NU	NU	NU	NU	NU	L-H	1168.0	138	4865743.4	4858718.9	7024.6
6th	NU	NU	NU	NU	NU	NU	NU	M-M	1047.2	139	4882861.6	4874989.9	7871.7
7th	NU	NU	NU	NU	NU	NU	NU	H-M	811.5	143	4968391.7	4958099.3	10292.4
8th	NU	NU	NU	NU	NU	NU	NU	M-H	811.5	143	4968391.7	4958099.3	10292.4
9th	NU	NU	NU	L-L	NU	NU	NU	L-L	686.0	147	5034340.5	5023438.7	10901.7
10th	NU	NU	NU	NU	NU	NU	L-L	L-L	658.9	139	4814844.7	4803999.6	10845.1

Note: SP stands for slaughter pigs, SOF for sows on farms, FPB for feral pig bleeding, RL for random lifestyle, RB1 for risk-based option 1, RB2 for risk-based option 2, FR for farmer reporting, HCR for human case reporting (HCR). H means either high intensity or high sensitivity, M either medium intensity or medium sensitivity and L either low intensity or low sensitivity; NU means not used; Ave B/C means average benefit cost ratio; Ave DS means average days saved; Ave GB means average gross benefit; Ave NB means average net benefit; Ave SC means average surveillance cost.

8.4.4. Portfolios with both early detection and high net benefit

This evaluation aims to balance speed of detection and net benefit. It does that by requiring the average days to detection to be within 5 days of the fastest detection time, then chooses portfolios with the highest net benefit. It achieves a detection time of slightly under 101 days, with an average net benefit $555,218 better than the fastest detection time. This is largely because surveillance cost is reduced by $563,011 on average. In comparison with the highest net benefit option, disease is detected on average 33 days earlier, but net benefit is reduced by $392,118 on average. Cost of surveillance is increased by $500,920 on average. These portfolios typically include the risk-based component RB2 and the classical surveillance components SP and RL as well as the low cost options FR and HCR (Table 8-10). This portfolio type is likely to appeal to policy makers as offering a moderate cost option with detection performance almost matching the much more expensive option of fastest detection.

Table 8-10. Top ten surveillance portfolios for Austeria in Australia with fast detection and high net benefit

Rank	SP	SOF	FPB	RL	RB1	RB2	FR	HCR	Ave NB	Ave DS	Ave SC
1st	H-M	NU	NU	H-M	NU	H-M	H-H	H-M	5403015.7	264.2	569779.2
2nd	H-M	NU	NU	H-M	NU	H-M	H-H	M-H	5403015.7	264.2	569779.2
3rd	H-M	NU	NU	H-H	NU	H-M	H-H	H-M	5402576.3	264.6	570280.0
4th	H-M	NU	NU	H-H	NU	H-M	H-H	M-H	5402576.3	264.6	570280.0
5th	H-M	NU	NU	H-M	NU	H-M	H-H	H-H	5401515.4	264.3	571279.5
6th	H-M	NU	NU	H-H	NU	H-M	H-H	H-L	5401223.5	264.2	571187.8
7th	H-M	NU	NU	H-H	NU	H-M	H-H	L-H	5401223.5	264.2	571187.8
8th	H-M	NU	NU	H-H	NU	H-M	H-H	H-H	5401081.5	264.6	571774.8
9th	H-M	NU	NU	H-H	NU	H-M	H-H	M-M	5400833.8	264.2	571577.5
10th	H-M	NU	NU	H-H	M-H	H-M	M-H	H-M	5394561.9	264.0	577701.2

Note: SP stands for slaughter pigs, SOF for sows on farms, FPB for feral pig bleeding, RL for random lifestyle, RB1 for risk-based option 1, RB2 for risk-based option 2, FR for farmer reporting, HCR for human case reporting (HCR). H means either high intensity or high sensitivity, M either medium intensity or medium sensitivity and L either low intensity or low sensitivity; NU means not used; Ave B/C means average benefit cost ratio; Ave DS means average days saved; Ave GB means average gross benefit; Ave NB means average net benefit; Ave SC means average surveillance cost.

8.4.5.Portfolios with speedy detection and high benefit cost ratio

Interestingly, when the requirement for early detection is specified (within 5 days of the fastest detection option), this decision rule can no longer select very cheap but poorly effective portfolios, and the results are almost indistinguishable from the previous option, except that RB1 is now chosen as well as RB2, because this boosts the benefit cost ratio and only increases the cost by a modest amount. Therefore policy makers could be presented with both results, and can decide whether to make the small additional investment in RB1. The top ten selections are presented in Table 8-11.

Table 8-11. Top ten surveillance portfolios for Austeria in Australia with fast detection and high benefit cost ratio

Rank	SP	SOF	FPB	RL	RB1	RB2	FR	HCR	Ave B/C	Ave DS	Ave GB	Ave NB	Ave SC
1st	M-M	NU	NU	H-H	M-H	H-M	H-H	H-M	16.4	264.2	5972373.6	5385756.9	586616.7
2nd	M-M	NU	NU	H-H	M-H	H-M	H-H	M-H	16.4	264.2	5972373.6	5385756.9	586616.7
3rd	M-M	NU	NU	H-H	M-H	H-M	H-H	H-H	16.3	264.2	5972373.6	5384256.9	588116.7
4th	H-M	NU	NU	H-H	M-H	H-M	M-H	H-M	16.2	264.0	5972263.2	5394561.9	577701.2
5th	H-M	NU	NU	H-H	M-H	H-M	M-H	M-H	16.2	264.0	5972263.2	5394561.9	577701.2
6th	H-M	NU	NU	H-H	M-H	H-M	M-H	H-H	16.2	264.1	5972263.2	5393059.2	579204.0
7th	H-L	NU	NU	H-H	M-H	H-M	H-H	H-H	16.1	264.3	5972400.6	5375916.0	596484.6
8th	H-L	NU	NU	H-H	M-H	H-M	H-H	M-H	16.1	264.3	5972400.6	5375916.0	596484.6
9th	H-L	NU	NU	H-H	M-H	H-M	H-H	H-H	16.0	264.3	5972400.6	5374418.8	597981.8
10th	H-M	NU	NU	H-H	M-H	H-M	H-M	H-M	16.0	264.0	5972263.2	5386894.3	585368.8

Note: SP stands for slaughter pigs, SOF for sows on farms, FPB for feral pig bleeding, RL for random lifestyle, RB1 for risk-based option 1, RB2 for risk-based option 2, FR for farmer reporting, HCR for human case reporting (HCR). H means either high intensity or high sensitivity, M either medium intensity or medium sensitivity and L either low intensity or low sensitivity; NU means not used; Ave B/C means average benefit cost ratio; Ave DS means average days saved; Ave GB means average gross benefit; Ave NB means average net benefit; Ave SC means average surveillance cost.

8.4.6. The final set of optimal surveillance portfolios

Depending on the criteria used by policy makers for decision making on surveillance, different optimal surveillance portfolios can be chosen as discussed in each section above. They are summarized in Table 8-12.

Table 8-12. A summary of the best surveillance portfolios for Austeria in Australia

Portfolio	SP	SOF	FPB	RL	RB1	RB2	FR	HCR	Ave DS	Ave B/C	Ave NB
Fastest	H-H	H-H	NU	H-H	H-H	H-H	H-H	H-H	269	8.5	4847797.5
BestAveNB	NU	NU	NU	H-H	NU	NU	M-H	H-L	232	132.9	5795134.1
BestAveB/C	NU	NU	NU	NU	NU	NU	NU	L-L	124	1859.2	4494843.6
FastAveNB	H-M	NU	NU	H-M	NU	H-M	H-H	H-M	264	15.6	5403015.7
FastAveB/C	M-M	NU	NU	H-H	M-H	H-M	H-H	H-M	264	16.4	5385756.9

Note: Fastest means the portfolio taking the least detection time. BestAveNB means the portfolio with the highest net benefit; BestAveB/C means the portfolio with the highest benefit cost ratio; FastAveB/C means the portfolio with the highest benefit ratio and taking 101 days to detect the disease; FastAveNB means the portfolio with the highest net benefit and taking 101 days to detect the disease. H means either high intensity or high sensitivity, M either medium intensity or medium sensitivity and L either low intensity or low sensitivity; NU means not used; Ave B/C means average benefit cost ratio; Ave DS means average days saved; Ave NB means average net benefit. SP stands for slaughter pigs, SOF for sows on farms, FPB for feral pig bleeding, RL for random lifestyle, RB1 for risk-based option 1, RB2 for risk-based option 2, FR for farmer reporting, HCR for human case reporting (HCR).

237

None of the optimal surveillance portfolios has included FPB, the approach of using hunters to undertake volunteer surveillance among feral pigs, by collecting blood samples from pigs they kill. This is because the intensity of surveillance that was considered feasible for them to conduct was too low for them to detect infection earlier than the on-farm options. Sampling sows on farms is costly and also relatively poorly effective, so was only included when cost of surveillance was not considered in the selection of portfolio components. All of the other surveillance components are included in some of the portfolios, depending on the weight given to speed of detection versus minimizing cost.

8.5. Discussion

8.5.1. Feasibility of optimizing zoonotic disease surveillance

There has been extensive discussion in the scientific literature and in national policy circles about achieving improved disease surveillance, especially for emerging diseases, through integrating different surveillance methods and adopting a risk-based approach. However there has been no standard procedure which countries could use to apply these principles to practical policy decisions.

This study took an example of a region with spatially defined populations of owned pigs, feral pigs, and associated people, and created a fictitious zoonotic disease that involved all three populations. The disease was then modelled on this spatial population structure using the HandiSpread disease model, adapted to represent the epidemiological scenario. By using 99 iterations of the model with each based on a separate random number seeding process, variability in the disease process was represented while keeping all parameter settings constant. The 99 models were then run with all 72 surveillance components and sub-components, to test how long each surveillance strategy took to detect the disease when evaluated on equal terms. This can only be achieved with a model such as HandiSpread, where each biological process within the model is independently seeded, and by using 99 different starting seeds, the same epidemics could be run repeatedly, while varying the surveillance strategies.

Therefore it was technically feasible to produce 99 replicates of an epidemic with biological variability represented, and to determine the detection success of each of 72 surveillance strategies which used a single component and single sub-component. Detection success was determined by measuring days from incursion to detection, number of pig herds infected by the date of detection, and proportion of outbreaks detected in less than 365 days. The number of people infected was also recorded, and the economic impact in the pig population and the human population was estimated. The cost of each surveillance component/sub-component combination was also measured. This data then allowed the costs and economic benefits of each surveillance strategy to be estimated for each of the 99 epidemic replicates.

The challenge was then to develop a procedure for logically combining individual surveillance procedures into portfolios which would on average detect disease outbreaks more rapidly and/or more cost-effectively than a single procedure alone. Even with only 8 components, each with 9 sub-components, there were 100 million portfolios to choose from. The OptiSurv procedure was therefore developed, which screens potential combinations of components/sub-components, excludes ones which are inferior, and narrows the focus to those which detect the disease most rapidly, and where appropriate provide the greatest economic benefit. Because of the large number of possible portfolios, many related portfolios produce results which differ to only a slight degree, so it is better to think of clusters of suitable or unsuitable portfolios.

It seems intuitively reasonable to expect economic measures alone to identify optimal portfolios, but in fact (at least under the conditions of this case study) if either net benefit or benefit-cost ratio alone was used as a decision rule, the top-ranked portfolios sacrificed considerable speed of disease detection for moderate increases in economic outcome, while using minimum time to detection alone as the decision rule process an extremely expensive portfolio for only a small reduction in days to detection. It was concluded that the optimal strategies for most situations were those which produced intervals from incursion to detection which were close to but not equal to the shortest detection time, and did so with high net benefit or benefit-cost ratio. Such portfolios were not optimal on either detection speed or economic performance, but they provided the best balance between the two. Either economic measure was equally suitable to select a portfolio. In special situations one of the other decision rules may be preferred, but in general the balanced portfolio would be preferred. Thus it did prove possible to develop a structured process which identified optimal surveillance portfolios comprising several surveillance components from what might seem initially to be a bewildering range of possibilities. Once the disease model is set up and the optimization task defined, the procedure can be easily updated and re-run when new data is available, making this tool useful for evaluating optimal surveillance portfolios for other diseases.

Estimating the economic impact of a disease in pigs requires considerable data, but is relatively straightforward. The main challenge in applying economic analysis to a zoonotic disease is how to produce an economic measure for the effect of a disease on human health, which can be combined with the economic impact data for animals. The use of disability-adjusted life years (DALYs) is now broadly accepted as a way of quantifying the severity of disease impacts in people, but it is a purely epidemiological measure, not an economic one. The issue of how to put an economic value on a DALY is much more controversial[17], and there is no consensus on whether it is acceptable to do so, and if so

---

[17] Health of Nations: The Value of a Statistical Life. Australian Safety and Compensation Council, Australian Government, July 2008 pp. 194.

how to do it. The problem in considering economic aspects of zoonoses is that a monetary value must be assigned to a DALY in order to undertake the analytical procedure we have used. We used the human capital approach for estimating the impact of a DALY due to Austeria. This approach was consistent with the methodology used for the estimation of economic impact of Austeria on pig production, and the impacts could then be combined. In one sense this treats human life as a commodity with a price, which is generally regarded as inappropriate. However another way of viewing it is that the person who suffers the effects of a disease does in fact suffer an adverse economic effect, and representing that specific effect is acceptable, while acknowledging that it is a clear under-valuing of the overall significance of a DALY. An alternative to the human capital approach which has been put forward is the Willingness To Pay (WTP) approach (Mishan, 1971; Viscusi & Aldy, 2003). However, using WTP to estimate the economic loss of both infected animals and humans would be technically challenging (Howe, Häsler, & Stärk et al., 2013; Narrod, Zinsstag, & Tiongco, 2012), and suffers from as many ethical concerns as human capital, so was not applied in this case study.

8.5.2. Comparison of a surveillance portfolio with a single surveillance component

Our study demonstrates that combining surveillance components into portfolios indeed advances the performance frontier in terms of technical and economic efficiency when compared with that of using single component strategies, evaluating these two alternatives on the basis of detection speed and achieving the highest net benefit. For instance, the risk-based surveillance component (RB2), used with high surveillance intensity and detection sensitivity, could achieve the fastest detection among all the individual surveillance components. On average, it would take 184 days to detect Austeria, however, the optimal surveillance portfolios only took 96 to101 days to detect the disease. Further, the optimal surveillance portfolio with highest net benefit would save 188,516 dollars compared to the most economically efficient single component (RL with high intensity and high sensitivity).

Table 8-13 Comparison of the optimal surveillance portfolios with single components for two different surveillance decision rules

Criterion	Optimal Portfolio	Best performing single component
Average DS	SP/SOF/RL/RB1/RB2/FR/HCR: 269 days	RB2: 181 days
Average NB	FR/HCR: $5,795,134	RL: $5,606,618

Note: DS means days saved, it equates to 365 minus the number of days for detection; NB means net benefit.

8.5.3. Practical consideration of different criteria for surveillance decision making

Our study demonstrates that there are many choices for policy makers for the selection of most suitable surveillance strategies. The idea is illustrated in Figure 8-3. A (in red) represents the surveillance option with the highest benefit cost ratio; B (in green) means the option with the highest net benefit (the

highest economic efficiency) and C (in orange) denotes the fastest detection of disease (the highest technical efficiency).
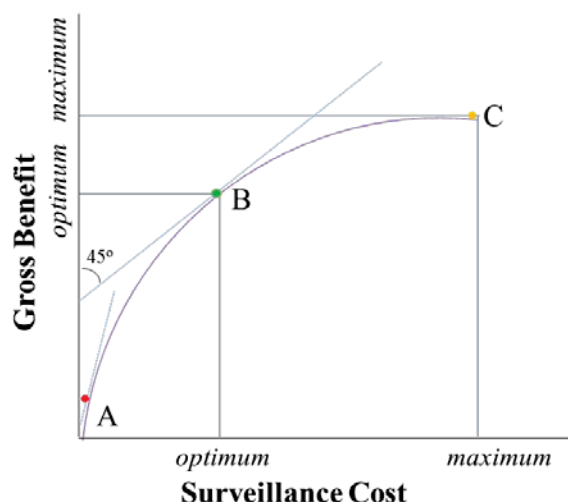

Figure 8-3. Optimal performance of a surveillance strategy
Note: adopted from Howe, K., B. Häsler, et al. (2013). "Economic principles for resource allocation decisions at national level to mitigate the effects of disease in farm animal populations." Epidemiology and infection **141**(01): 91-101.

8.5.3.1. Maximizing economic efficiency

Resources are always scarce and have opportunity costs. For disease prevention and control, resources have to be allocated among different disease management programs; for the same disease, resources also need to be allocated between surveillance and disease mitigation. Hence, striking the right balance in resource allocation between disease detection and mitigation may mean making what appears to be a suboptimal investment in surveillance, in order to be sure of having sufficient resources to undertake mitigation if the surveillance discovers that disease is present. Therefore it may be economically rational to reduce the investment in surveillance to a moderately suboptimal level, even if resources are available. The "balanced" portfolios which require both days to detection and economic benefit to be mildly suboptimal are likely in most cases to be the preferred strategy. In case of Austeria, Figure 8-4 shows that for fairly early detection, a minor delay has only a small effect on the number of farms which become infected, whereas at a later stage a delay of equal length is much more serious.
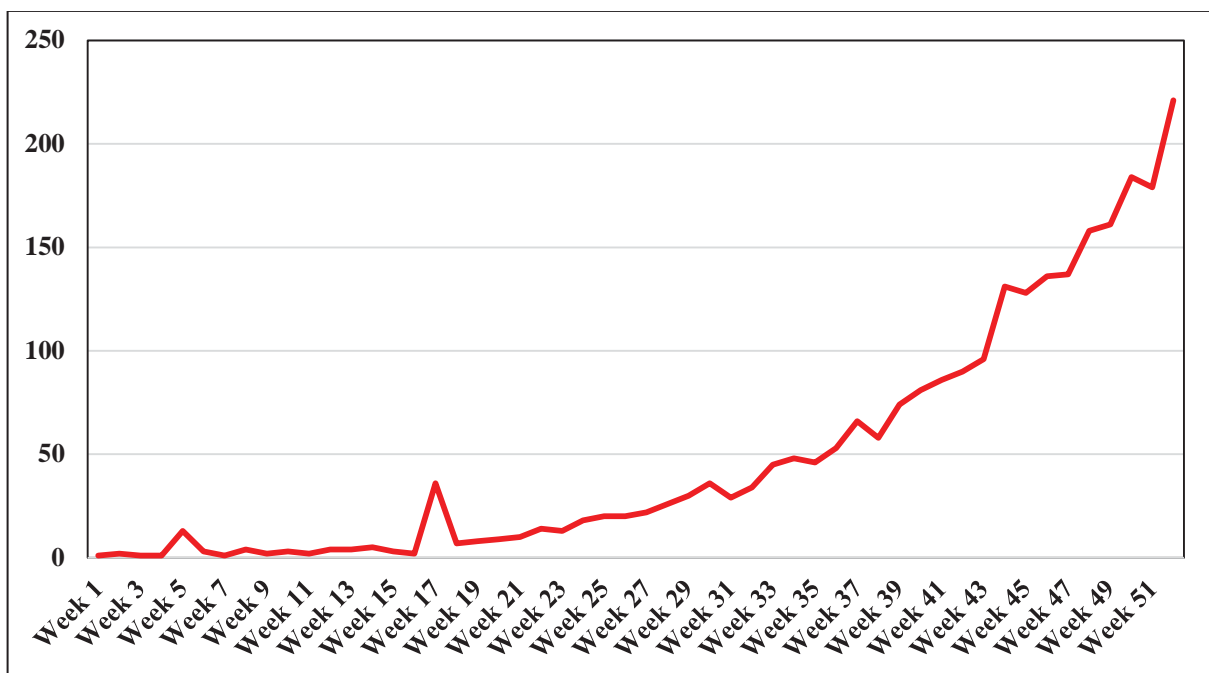
Figure 8-4. The weekly count of infected farms caused by a medium sized Austeria outbreak in Queensland, Australia

8.5.3.2. Maximizing technical efficiency

For diseases, such as Ebola haemorrhagic fever, that can spread easily and rapidly cause high case fatality rates among people in a short period of time, a country may choose to adopt a portfolio which minimizes time to detection, despite the high cost.

8.5.3.3. Further development of economic analysis of zoonoses

The economic analysis procedure used in this study was specific to this disease, but represents one of the examples which is being used to guide development of a generic form of economic analysis of zoonoses, although the work is outside the scope of this project.

A prototype of a generic form of analysis in spreadsheet form is shown below in Figure 8-5 to illustrate the next stage of development currently underway from work undertaken in this and other related economic studies of specific diseases.

Figure 8-5 Generic form of economic analysis of zoonoses being developed -beyond the scope of this project

8.6. Conclusions

The study introduced a practical way for screening all feasible surveillance portfolios comprising one or more surveillance components, and using decision rules based on one of more of the criteria speed of detection, cost of surveillance and net benefit of surveillance to identify a cluster of surveillance portfolios, any one of which can best satisfy the chosen decision rule. It demonstrated that under the circumstances of the case study multi-component surveillance strategies perform better than single surveillance component both epidemiologically and economically. In general, a surveillance portfolio which is mildly suboptimal for both days to detection and net benefit of surveillance is likely to be preferred by policy makers, because it achieves the best balance between the two objectives.

The approach employed and the OptiSurv tool used in the study are generalizable for optimization of surveillance programs for other zoonotic diseases.

8.7. References

Guo, X., G. Claassen, et al. (2014). "A conceptual framework for economic optimization of single hazard surveillance in livestock production chains." Preventive veterinary medicine **114**(3): 188-200.

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

Holtkamp, D., J. Kliebenstein, et al. (2007). Assessment of the economic impact of porcine reproductive and respiratory syndrome virus on United States pork producers.

Howe, K., B. Häsler, et al. (2013). "Economic principles for resource allocation decisions at national level to mitigate the effects of disease in farm animal populations." Epidemiology and infection **141**(01): 91-101.

Hutwagner, L., T. Browne, et al. (2005). "Comparing aberration detection methods with simulated data." Emerg Infect Dis **11**(2): 314-316.

Mishan, E. J. (1971). "Evaluation of life and limb: a theoretical approach." The Journal of Political Economy: 687-705.

Murray, C. J. and A. K. Acharya (1997). "Understanding DALYs." Journal of health economics **16**(6): 703-730.

Narrod, C., J. Zinsstag, et al. (2012). "A one health framework for estimating the economic costs of zoonotic diseases on society." EcoHealth **9**(2): 150-162.

Polgreen, P. M., Z. Chen, et al. (2009). "Optimizing influenza sentinel surveillance at the state level." American journal of epidemiology: 1300-1306.

Prattley, D., R. Morris, et al. (2007). "Application of portfolio theory to risk-based allocation of surveillance resources in animal populations." Preventive veterinary medicine **81**(1): 56-69.

Scarpino, S. V., N. B. Dimitrov, et al. (2012). "Optimizing provider recruitment for influenza surveillance networks." PLoS Comput Biol **8**(4): e1002472.

Sparks, R. (2013). "Challenges in designing a disease surveillance plan: What we have and what we need?" IIE Transactions on Healthcare Systems Engineering **3**(3): 181-192.

Viscusi, W. K. and J. E. Aldy (2003). "The value of a statistical life: a critical review of market estimates throughout the world." Journal of risk and uncertainty **27**(1): 5-76.

Wang, X., D. Zeng, et al. (2010). "Comparing early outbreak detection algorithms based on their optimized parameter values." Journal of Biomedical Informatics **43**(1): 97-103.

Zhang, L., I. Jamal, et al. (2011). "Optimization of event detection methods for disease surveillance." Epidemiologie et Sante Animale(59/60): 196-198.

9. General discussion

9.1. Goal of the thesis

The thesis was focused on development, application and evaluation of a software toolbox, HandiResponse, for identifying the most appropriate surveillance portfolio for zoonotic diseases, especially emerging diseases, taking account of the influence of epidemiological risk factors and environmental influences on the spatial distribution of the risk of occurrence of the disease of concern.

The major contributions of the author in the project include (i) development of the concept and design of the architecture of HandiResponse; (ii) specifying objectives, expected outputs and design of detailed steps of the technical process that generate the outputs, identification of techniques employed in each module; (iii) parameterization of disease spread models for HPAI H5N1 and Austeria based upon literature review and consultations with topic experts and (iv) testing the proto-type version(s) of HandiResponse and providing feedback for further refinement.

9.2. Studies undertaken in the thesis

The project discussed in this thesis is about designing and testing a set of tools that could enable an epidemiologist in any country to use available evidence about an emerging disease which might enter the country to develop risk landscape for the disease and use such knowledge to plan an optimal mix of surveillance activities to detect the disease promptly and in a cost-effectively manner if it enters the country. Just as the thesis is being submitted, there is worldwide concern about spread of Zika virus, which is the latest example of the series of emerging diseases that have spread around the world in recent years, and for which HandiResponse could provide countries with a valuable tool.

The structure of the thesis is illustrated by the diagram below. It forms part of a wider project to develop a comprehensive planning and management strategy for emerging diseases.

HandiResponse starts from risk assessment on possible disease occurrence, then guides the user through the steps of risk mapping, modelling disease spread, evaluating the potential development of a surveillance portfolio taking account of spatial variation in disease risk, and finally portfolio optimization. All the processes are facilitated by different modules within HandiResponse: HandiMap, HandiSpread, HandiSurv, HandiEcon and OptiSurv respectively.
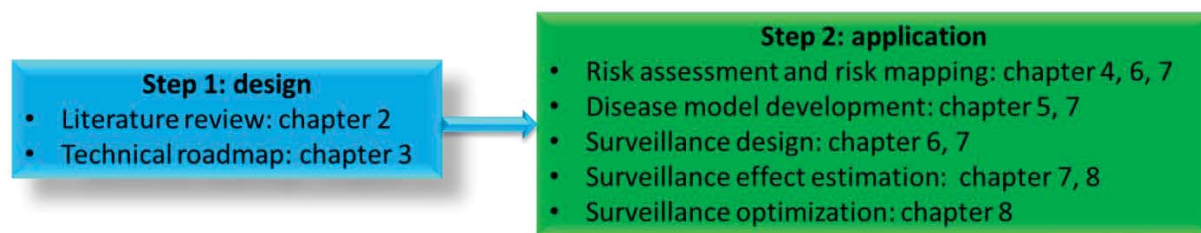
Figure 9-1 Thesis structure

Emerging and re-emerging infectious diseases (ERID) are capable of generating sizable economic impact and can cause social instability (Brahmbhatt, 2006; Narrod, Zinsstag, & Tiongco, 2012; UNDG, 2015; WorldBank, 2014). To prevent and mitigate the negative impacts of ERID, it is imperative to have a sensitive surveillance strategy for early disease detection. Furthermore, from the economic perspective, resources are always scarce and have opportunity cost, so investment in a surveillance program has to demonstrate that it can maximize the utility function of allocated resources.

Two notable development in disease surveillance system design in recent years deserve highlight: firstly, risk-based surveillance has gained in popularity and represents an important advance for modern disease surveillance program development (Stark, Regula et al., 2006). Risk-based surveillance programs can outperform the traditional non-risk-based (Alba, Casal et al., 2010; Hadorn, Rufenacht et al., 2002; Kahn, 2006; Kuiken, Leighton et al., 2005; Kulasekera, Kramer et al., 2001; Kulldorff, Fang, & Walsh, 2003; Presi, Staerk et al., 2008; Reist, Jemmi, & Staerk et al., 2012; Schwermer, Reding, & Hadorn et al., 2009; Tavornpanich, Gardner et al., 2006; Walsh & Miller, 2010; Willeberg, Nielsen, & Salman, 2012). Secondly, a surveillance program rarely has only one component, rather it is often a network of networks or a system comprising different components varying in surveillance subject, location, season, event, diagnostic method etc. (Eidson, Kramer et al., 2001; Heymann & Rodier, 1998; Morse, 2007; Mostashari, Bunning et al., 2001; Mostashari, Kulldorff et al., 2003).

It has however proved challenging in many situations to plan an integrated risk-based approach for surveillance and optimize a portfolio comprising several different surveillance approaches because of policy constraints, difficulty in getting access to all the required data, inadequate epidemiological capacity and resources to design and manage the program and other difficulties (Alba, Casal et al., 2010; Cameron, 2012; Carroll, Au et al., 2014; Christensen, El Allaki, & Vallières et al., 2014; adorn & Stark, 2008; Martin , Cameron et al., 2007a; Ortiz-Pelaez, Pfeiffer et al., 2006; Paul, Held, & Toschke, 2008; Stark, Regula et al., 2006; Tsai, Scott et al., 2009; Wieland, Brownstein et al., 2007), These challenges are usually more prominent in developing countries. Hence, most of the risk-based

surveillance activities have been so far implemented in developed economies, whereas they could provide greater benefit if applied in resource-constrained countries.

HandiResponse was developed to contribute to dealing with these challenges. The specific objectives of HandiResponse are (i) to visualize the disease risk landscape and identify hotspots where the infectious disease under study is most likely to occur if it becomes established in the country; (ii) to evaluate economic benefit and costs of each surveillance option and their combined effects as components of a portfolio; (iii) to define optimal use of resources for a surveillance program through selection of an epidemiologically effective and/or an economically efficient surveillance portfolio across predefined risk categories such as geographic area, species, sectors, types of stakeholders, etc. It comprises four modules: HandiMap, HandiSurv, HandiEcon and OptiSurv. Besides, the program needs to be used in conjunction with HandiSpread, a disease simulation modelling program (Figure 9-2). The purposes and the outputs of these models are summarized in Table 9-1.
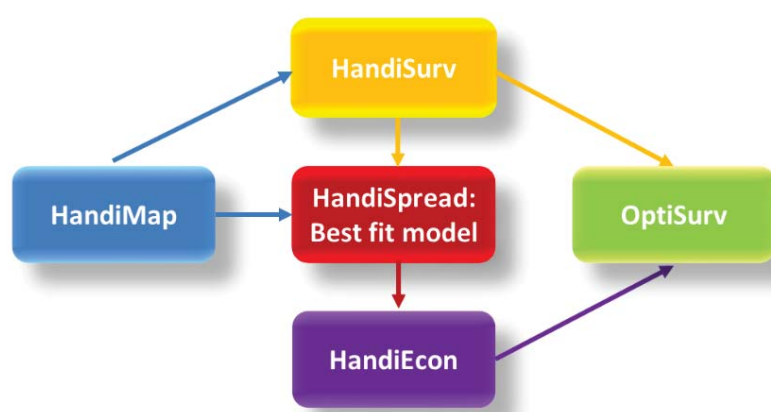
Figure 9-2. Structure of HandiResponse and linkage between different modules

Table 9-1. The purpose and output of different modules of HandiResponse

Purpose	Output
HandiMap: development of risk landscape for an infectious disease	• risk map(s) that display the "height" of the risk at each location. The maps can be either in kernel smoothed format or raster format • a risk score file which transfers the "risk height' information to HandiSpread, for adjusting the susceptibility of particular locations to occurrence of the disease, and hence allows HandiSpread to take account of risk level in representing transmission of the disease
HandiSurv: defining all the possible surveillance components and subcomponents for an infectious disease	• a matrix summarizing surveillance subject, location, modality, intensity, sensitivity and cost
HandiEcon: estimation of economic effect of a disease or a given surveillance approach	• a set of MS Excel spreadsheets summarizing (i) economic effect of an uncontrolled infectious disease by time, location and species and (ii) economic effect of an infectious disease under a given surveillance option
OptiSurv: optimization of surveillance approaches according to the predefined criteria	• the optimal surveillance portfolio(s) meeting a given criterion

Although as described in the literature review and later chapters there have been many papers published both on methods of undertaking surveillance for emerging diseases and on specific investigations of particular diseases, there has been no previously published step by step generic system for a country to use in planning and managing how it can use the techniques and the scientific information in the national situation.

9.3. Classification of diseases

Making the system generic both with respect to diseases which can be considered was the first challenge. It is clearly not possible to have a system matched both to all known diseases and to previously unknown emerging diseases! The approach adopted was to develop a structure which allowed diseases to be categorized into a limited number of groups, which have been termed "epitypes" for the purpose of the thesis. The classification system described in Chapter 3 was adapted from a World Health Organization report on zoonoses, and allows the development of an analytical approach matched to the specific epidemiological characteristics of each epitype, since the diseases classified

within each epitype share most of their core features in common, even though they may appear different in their manifestations. An assessment of zoonoses did not identify any that fell outside the overall classification structure. By far the largest number of zoonoses fit into the epitype of direct zoonoses, so this group has been divided into several sub-epitypes, since there are important epidemiological differences, for example, between respiratory route transmission and foodborne transmission, even though they are both direct. This thesis has concentrated principally on developing the later stages of the analytical process using examples of direct zoonoses, and further work will be required in future to design and build a module of HandiResponse which assists the user to decide which epitype and sub-epitype a disease belongs to, and then to record defining epidemiological characteristics of the disease, so that the mapping stage can proceed smoothly.

9.4. Steps in the HandiResponse surveillance planning process

As illustrated in Figure 9-2, an analysis using HandiResponse typically starts with HandiMap, which produces risk map(s) and generates a risk file with a score for each epidemiological unit. Step two, the information on risk scores is fed into HandiSpread for the development of a model representing the spread of a newly introduced disease temporally and spatially. Step three, the predicted temporal and spatial distribution of a disease informs the design of different surveillance options including risk-based ones for it; then, the model in HandiSpread will be used to simulate the potential impacts of each surveillance option. Step four, the metrics, such as number of farms and/or persons infected by the day of detection, from the simulation of surveillance options are input into HandiEcon for estimating the economic effect of each surveillance option, in contrast to the economic impact of an undiagnosed disease outbreak. Step five, the statistics produced by HandiSurv and HandiEcon are used by OptiSurv to generate optimal surveillance portfolios.

9.5. Development of a risk map

Chapter 4 demonstrated how to use HandiMap to generate disease risk maps in a stepwise approach. As a major method of spatial epidemiology, disease risk mapping can help test a hypothesis, provide a direction for surveillance and control efforts or evaluate the actual or potential effectiveness of an intervention (Brownstein, Freifeld, & Madoff, 2009; Kitron, 2000; Tatem, Adamo et al., 2010). It can also provide evidence for substantiating progress towards global health commitments (Pigott, Howes et al., 2015). However, a recent review revealed that out of 176 globally important infectious diseases with a strong rationale for mapping, only 4% have been adequately mapped (Hay, Battle et al., 2013).

A major reason why scientific data on disease risk has not commonly been converted into a risk map is the effort normally required to bring together the data sets necessary to draw a risk map, and integrate the evidence into a digital map which reflects the available evidence. HandiMap now provides a tool for

doing this across the full spectrum of diseases where spatial representation is helpful. To do this requires access to both country geography and spatial data on putative risk factors and proxy variables which can be overlaid on the country geography. The first step of providing country geography is simple in principle, but there were challenges to resolve along the way in providing the data to HandiMap. Global digital maps are available which provide shape files for all countries and for the main subdivisions within each country, although subdivisions change and digital maps may have outdated boundaries for some subdivisions. Shape files for lower level administrative areas are not generally available at global level, so must be obtained from national sources. For the study of Mongolia in Chapter 5, boundaries of provinces (known as aimags) were available on global map data but boundaries of districts (known as soums) were required since sampling for CCHF was undertaken at this level, and these were obtained after some effort. There is no publicly available set of shape files for the lowest administrative area in Mongolia, sub-district or bag, and these shape files were only available through direct negotiation with the official who had produced them. However they were not necessary for this particular study. Mongolia illustrates another practical problem resulting from the large size of the country, its very large east-west extent, and its high latitude (which makes effects of the curvature of the earth very pronounced). The standard digital map for Mongolia used three separate somewhat incompatible map projections for different parts of the country, and it required expert help to create a single projection with smooth unbroken connections between the parts. However in working with data for various countries we have found that any problems with administrative boundary maps can be solved and once the map for a country is correctly established in the data repository, it can be used as much as needed. The other layers such as Google hybrid can also be overlaid on the administrative boundary map.

The development in recent years which has made disease risk mapping much easier and potentially much more informative has been the rapid expansion in the range of global remote-sensed satellite images which are now available. Until a decade or so ago, satellite images were quite restricted in the types of data available, and costs of obtaining digital images were high. However the range of data sets of relevance to disease control has increased greatly over the last decade, and they are now mostly available at no cost from web sites. The resolution has also improved considerably, and the interval between successive images being produced and made available has reduced. This makes disease risk mapping a very straightforward and informative way of exploring disease distribution. Since HandiMap can process any digital raster image provided to it from satellite data, it is a very effective and straightforward way to build risk maps for a disease of concern if remote-sensed data for risk factors and/or proxies is available.

In Chapter 5, HPAIV H5N1 outbreak in southern Vietnam was used as an example to demonstrate the procedures used to create a risk map in HandiMap, and to compare maps which use different risk factors.

252

Compared with other widely available mapping tools, HandiMap has the following merits: (i) at least one disease risk map under each epitype has been developed or will be developed for program users to learn by transferring the methods from the demonstration disease to the disease of interest to them; (ii) more than 30 global environmental and climatic descriptor layers have been stored in the IRIS data server and can be accessed freely, so an epidemiologist can explore the use of different layers without needing to download them from multiple servers and install them in a standard GIS package. The number of layers stored in the repository will continue to increase, and the file sizes are very large so it is much better to have them stored on a single site than to have individual users store them on local hard drives. The layers will also need to be regularly updated as new releases become available, and this is best done through a single server site; (iii) risk values by epidemiological unit can be exported and used by other modules of HandiResponse (or other disease modelling programs) to simulate disease spread, and assess the effect of any disease surveillance and mitigation strategies and (iv) the tool is free of charge.

HandiMap is a convenient generic tool for estimating and visualizing the risk of occurrence for a disease of interest. It can facilitate effective communication on infectious diseases and support decision making for surveillance and disease interventions, and it offers particular benefits in resource and data-sparse environments. There have been numerous papers published on factors influencing the spatial distribution of avian influenza H5N1 outbreaks, especially for Vietnam, because one of the most comprehensive spatial data sets from this global disease event was the one for Vietnam. It is possible to produce risk maps for Vietnam based on the risk factor information, but the question which has not previously been answered is whether the risk map is genuinely informative about where outbreaks are likely to occur, on a fine enough scale to support surveillance and control activities. Therefore a review of all available papers was undertaken, and the findings used to formulate two initial risk maps for southern Vietnam, one based entirely on environmental variables derived from satellite data and considered by some investigators to influence disease occurrence, and one using poultry and human population data together with selected environmental factors. Then a third map was produced, in which most of the factors in the earlier two maps were included, plus variables which were proxies for ease of movement of poultry around the region, since no direct data was available on movement patterns.

The issue which then had to be dealt with was how to test whether the risk maps had epidemiological validity. Therefore to test the value of the risk maps, a simulation model of avian influenza H5N1 (termed the base model because it did not make use of a risk map) was constructed in the modelling program HandiSpread, and parameters adjusted so that the model gave a cumulative temporal epidemic curve not significantly different from the actual epidemic curve, by the Kolmogorov-Smirnov statistic. Disease models have become valuable tools that help to provide insights on disease spread, generate testable hypotheses, predict disease outbreaks and assess the effect of surveillance and disease mitigation measures (Barnabas, Laukkanen et al., 2006; Brisson, Edmunds et al., 2000; Keeling,

Woolhouse et al., 2001; Koopman, 2005; Merler, Ajelli et al., 2015; Mossong, Hens et al., 2008; Myers, Rogers et al., 2000; Riley, 2007; Woolhouse, 2011). Spatial and temporal modelling offers a way to make predictions about where and how fast a disease would spread. Such pieces of information are particularly important at early stages of emerging infectious outbreaks when little knowledge has accumulated about disease behaviours. However, there remains a need to further develop and test models of emerging diseases and to build them into decision support systems (Bettencourt & Ribeiro, 2008; Kao, 2002; Lawson & Leimich, 2000). HandiResponse offers an answer to these needs by providing risk mapping, spatio-temporal modelling and other components of a decision support system for detecting disease incursions.

Very few examples have been reported in which model predictions for an emerging disease have been compared with field evidence of the spatio-temporal spread of the same disease. In Chapter 5 the benefit was evaluated of modelling on a risk landscape generated in HandiMap when compared with the base model – the same disease model without the adjustment in spread probability provided by the risk map. To achieve this, the first step was to develop the base model in HandiSpread of avian influenza H5N1 using research data and expert opinion, and to compare the temporal and spatial pattern with field data from the 2004-5 second epidemic wave in southern Vietnam. This first model gave good temporal fit to the real epidemic wave, but spatial fit was only moderate. Two different variants of this base model were chosen, which had mildly different settings for one set of spread parameters, but both of which gave good temporal fit to the actual epidemic wave. The field data set covering the second epidemic wave used for this comparison was chosen because it was relatively free from biases due either to under-reporting or poultry vaccination.

Having produced the three different risk maps (E, P and M), the next issue was to assess what adjustment to infection susceptibility was appropriate to use for each of the five risk levels allocated to communes in the risk maps. The key issues were whether to use linear or escalating effect of increasing risk level, and what percentage change in susceptibility to use between different risk levels. To test these issues, a matrix of 162 adjustment factor combinations was constructed, and 99 replicate simulations were run for each combination. The modelled outbreaks were then tested for their temporal fit in comparison with the actual HPAIV H5N1 epidemic wave in southern Vietnam by Kolmogorov-Smirnov test and for spatial fit by calculation of the area under the Receiver-Operating Characteristic curve (AUC value). Out of 162 models tested, five candidates were selected as having the best fit of the risk adjusted models. The best one was the M-map m-r-3-10 linked to base model 4. This meant that the adjustment factor was escalating rather than linear, the reference level was the middle risk value 3, and the adjustment was 10%. The robustness of fit of the m4-m-r-3-10 model was tested by changing the relative importance of the movement risk component to the environmental descriptor component. The results showed that m4-m-r-3-10 still outperformed all other models.

The exercise revealed that M-model, using a combination of different risk factors and environment predictors, was the best option for representing HPAIV H5N1 outbreak in southern Vietnam between late 2004 and early 2005. In this model, poultry movement was hypothesized as the most important risk factor, followed by rice paddy coverage. Both risk factors have been confirmed to be responsible for HPAIV H5N1 outbreaks in South East Asia (Cao, Xu et al., 2010; Fang, de Vlas et al., 2008; Gilbert, Xiao et al., 2008; Martin, Pfeiffer, et al., 2011; Peterson & Williams, 2008; Pfeiffer, Minh et al., 2007; Rivas, Chowell et al., 2010; Tiensin, Chaitaweesup et al., 2005; Williams & Peterson, 2009; Xiao, Boles et al., 2005; Yupiana, de Vlas et al., 2010). In particular, poultry movement was confirmed to be significantly associated with the spread of HPAIV H5N1 outbreaks in Vietnam (Pfeiffer, Minh et al., 2007). E-model would be the second choice since the model was mainly defined by distribution of rice paddy. The least satisfactory model was P-model, which used only poultry and human population-based risk factors, with no environmental or movement factors included.

The disease spread model could also generate insightful epidemiological metrics to be used for informing surveillance and disease mitigation design. For instance, under m4-m-r-3-10 model, live bird markets and itinerant duck movement were responsible for 49 percent and 40 percent of HPAIV H5N1 infections respectively, and other forms of spread were minor. Efforts for surveillance and control should therefore give special emphasis to these forms of spread.

While this study was valuable in testing the value of a risk map in predicting disease spread, there are various issues which could not be fully resolved. The first is that the field data set used for comparison is itself known from experience of the Vietnam epidemic to be imperfect, although the best available (Gilbert, Xiao et al., 2008). If better data on H5N1 infection prevalence and clinical disease were available, the technical performance of models could possibly be better. Despite the large scale of the avian influenza H5N1 global epidemic, the amount of credible clinical outbreak data with spatial locations which was collected was very small, and the amount of data on infection prevalence even smaller. Far too little effort in major disease outbreaks of all kinds goes into recording and analyzing the spatio-temporal characteristics of the epidemic, and therefore there is very limited opportunity to improve the response capability for future events. HandiResponse and its underlying database HandiView provide a valuable tool for improving disease records, if they can be adopted by countries around the world through an international effort. The Ebola epidemic and now the spread of Zika virus around the world may provide the required stimulus for this to happen.

The second issue is that while it appears that environmental and movement risk factors were the most informative in guiding risk map development, the conclusions for this disease in this environment are at best tentative, and based on retrospective analysis rather than testing whether predictions of disease occurrence based on a risk map prove useful in surveillance.

9.6. Using risk mapping to inform surveillance

Therefore the next exercise in the project was to test the predictive ability of risk mapping. In Chapter 6, HandiMap was used to visualize the risk landscape for Crimean Congo Haemorrhagic Fever in Mongolia, and then to undertake surveillance to determine whether a disease incursion had occurred. The disease has been reported among human and livestock in neighboring parts of China and other countries such as Russia, Tajikistan and Kazakhstan in close proximity to Mongolia (Atkinson, Chamberlain et al., 2013; Hoogstraal, 1979; Yashina, Petrova et al., 2003; Yen, Kong et al., 1985). However neither animal infection nor human cases of CCHF had ever been reported in Mongolia. Livestock in Mongolia represents an essential basis for sustainable development, and conservation of the livelihood tradition of Mongolian herders. In 2010, more than 90 percent of people employed in the agricultural sector of Mongolia were engaged in animal husbandry. The traditional grazing practices can expose herders and domestic animals to tick bites. If ticks are infected with CCHFV, human infection would be inevitable. Besides, the growing mining industry in the country could increase the opportunities for susceptible miners working in the field to be exposed to ticks and hence to possible infection. The disease might exist in the country undetected because of low awareness and very limited disease-specific expertise.

A risk assessment on CCHF conducted in conjunction with this study suggested that most parts of Mongolia were suitable for *Hyalomma* tick presence, and there existed areas of Mongolia where *Hyalomma* ticks and other suitable ixodid species could complete their lifecycle in small and large mammals (Atkinson et al., 2013; Roger Hewson, 2013). Informed by a review of the published literature, the following environmental predictors were selected for CCHF risk map development: (i) proportion of land covered by shrub and grass, (ii) mean annual NDVI, (iii) mean annual precipitation and (iv) maximum temperature. Separate risk maps were developed based on predicted *Hyalomma* tick distribution and on CCHF distribution. They showed broad agreement on where high risk and low risk areas of the country are located, although the risk map for CCHF indicated a more extensive risky area than did the tick map, because of the different risk factors used. It is likely that the tick map is a more precise predictor, because the disease cannot persist long term in an area unless the host ticks are present. Targeted cross sectional serological surveys of people in herder families and their sheep were conducted in twenty one districts in the area of southern Mongolia identified as high risk on both types of maps, and in one district in the low risk northern part of the country. The survey showed for the first time that serological evidence of exposure of both sheep and people to CCHFV was present in all the high risk part of Mongolia evaluated, but there was no evidence of infection in either species in the district identified as low risk. Therefore the risk map was predictive of the localized presence of an important disease in a country previously considered to be free of it.

9.7. Development of optimal surveillance portfolios

The next step was to use the components of HandiResponse to test the effectiveness of risk-based and classical surveillance techniques for detecting a zoonotic disease incursion, and to assess the costs of surveillance and the benefits of prompt detection. This required comprehensive information on the spatial distribution of animal populations, and on economic aspects of the disease in at-risk animal and human populations. No suitable sets of required data could be identified for an Asian country, but data sets were available on spatial distribution of domestic pig populations and economic aspects of pig production for Australia, arising from an earlier study of a possible incursion of porcine reproductive and respiratory syndrome into Australia, which is free of the disease.

In Chapter 7, this information was used as the basis for inventing a zoonotic disease called "Austeria" as an example to test the effects of different surveillance options in disease detection. Australia has a large population of feral pigs which would undoubtedly be involved in spreading a disease of this nature, and it was essential that they be included in the evaluation of surveillance options. A risk map had already been developed for predicted density of feral pigs in all parts of Australia by national experts in the field of feral pig ecology. This map was therefore used together with ecological information from Australian studies to produce a map of home ranges for feral pig families throughout the country that reflected local densities, and each feral pig family was given a point location in its home range.

A temporal and spatial dynamic model was then developed to simulate the spread of Austeria over a period of 365 days following an incursion into a "backyard" or non-commercial pig herd in the State of Queensland, which has the largest number of domestic pig herds, plus a large population of feral pigs with varying density of families depending on local habitats.

Eight potential surveillance components were then identified, with each designed to detect the disease by focusing on different strata of at-risk pig populations or human populations, and different investigational strategies. Some components were risk-based active surveillance precisely targeted based on the risk map for feral pigs, others were classical surveillance sampling strategies, one component targeted feral pigs through voluntary sample collection by hunters, and others were passive disease reporting. Costs varied considerably according to the sample collection effort required. Each component had nine variants termed sub-components that differed in investigational intensity and detection sensitivity, making a total of 72 different options to be compared with the undetected outbreak. Using 99 iterations of the Austeria model per component/subcomponent combination, the technical efficiency of each surveillance approach was measured by (i) number of days for outbreak detection, (ii) number of simulated outbreaks detected, and (iii) number of farms infected by the time of detection, compared with the undetected outbreak.

The surveillance approaches fell into three groups according to the above mentioned three technical efficiency measurements. Risk-based option 2 (RB2) stood out as the most technically efficient approach and the feral pig blood sampling by hunters (FPB) was the least useful. The remaining six surveillance components fell in between and did not differ substantially in technical performance. The RB2 risk-based component with high intensity and high sensitivity detected the outbreak at a median of 108 days, and successfully detected over 93 percent of the 99 simulated outbreaks. At the time of detection, a median number of only 23 farms were infected. On the contrary, the least effective FPB sub-component with low intensity and low sensitivity could only detected the disease at a median number of 281 days after the disease was introduced, detected only 3 percent of the 99 simulated outbreaks, and when the disease was detected, a median number of 2, 515 farms were already infected.

The study demonstrated practically how a spatial and temporal model could help in designing surveillance strategies and assessing their technical performance. It revealed that it would take more than 4 months for the most efficient surveillance approach to detect a newly introduced slowly spreading disease like Austeria. Furthermore, it showed that the risk-based surveillance strategy RB1 which was expected to achieve the most rapid detection because it was focused on the highest feral pig density areas was in fact poorly successful because very few domestic pig herds were located in these areas, and they were the target population for surveillance. The study identified a new way in measuring a surveillance component sensitivity: how many outbreaks could be detected by a particular method out of all the simulated outbreaks. We concluded that all the three technical efficiency measurements were policy-relevant and should be used jointly in practical evaluation of surveillance options.

The study described in Chapter 7 left two important questions to be answered: (i) what would be the technical efficiency if surveillance components are used in combination as a portfolio rather than individually and (ii) what would be the economic efficiency of different surveillance portfolios.

Chapter 8 deals with these two issues. It does so through the development and application of OptiSurv, a surveillance optimization program, because the selection of an optimal portfolio, even from the restricted range of 72 different component/subcomponent combinations, requires a precisely structured evaluation procedure, and a goal-seeking strategy to select and rank the portfolios which best meet a predefined criterion, because the choice of criterion depends on the disease situation, and there are 100 million portfolios to choose from in this particular case. Resources are always scarce and have opportunity cost. Hence, decision making on disease surveillance strategies needs to be based upon sound technical and economic principles. Optimization of surveillance program for a disease or a hazard has been defined as maximizing surveillance performance within given or expected budget constraints (Guo, Claassen et al., 2014; Prattley, Morris et al., 2007), or it implicates maximizing cost saved for surveillance program without compromising or even increased technical performance (Hadorn & Stark, 2008; Wang Zeng et al., 2010). Superior technical performance means earlier

detection of a disease or a hazard, improved sensitivity and specificity (Hutwagner, Browne et al., 2005). A number of approaches have been employed to optimize the technical and in some cases economic performance of surveillance programs (Hadorn & Stark, 2008; Polgreen, Chen et al., 2009; Prattley, Morris et al., 2007; Scarpino, Dimitrov, & Meyers, 2012; Sparks, 2013; Zhang, Jamal et al., 2011). The approach adopted here builds on the earlier work of Prattley *et al.* to offer a standardised method for producing lists of optimal portfolios for each of the decision rules proposed in Chapter 8.

To achieve this, HandiSpread was used to construct a spatial and temporal model representing the spread of Austeria, as a test for a zoonotic disease, and to provide critical epidemiological metrics of the outbreak. The model was also used to estimate the potential effects of proposed surveillance options, in term of when the disease was detected, how many pig herds and how many persons would be infected as well as how many of the 99 simulated outbreaks could be detected. HandiEcon was then used to estimate the economic impacts of the disease: the epidemiological effect of Austeria on the pig industry in Queensland was monetized by employing a budgeting model, while disease burden (mortality and morbidity) caused by Austeria in people was estimated by the use of disability-adjusted life years and converted into an economic value by using the human capital approach. Outputs from HandiSpread and HandiEcon were then fed into OptiSurv to evaluate all possible portfolios and identify clusters of optimal surveillance portfolios. The version of HandiEcon used was one of a number of separate economic analysis procedures developed for different individual diseases and animal production systems, as preparatory work for the development of a generic economic analysis system for zoonoses, which is being produced outside the scope of this project.

OptiSurv could identify the optimal surveillance portfolio with regard to different decision rules. The portfolio with the shortest detection time was straightforward, and took an average of 96 days to detect the disease, across 99 replicates. In principle it would seem that the portfolio with the highest net benefit, or the one with the highest benefit cost ratio, would be the preferred policy option. However at least in the Austeria case study situation, the portfolio with the highest benefit/ cost ratio was simply the lowest cost one, with a very long delay to detection and the overall poor performance. Although the portfolio with the highest net benefit was better in terms of detecting the outbreak earlier but also unlikely to be the preferred policy option because of excessive delay in detection. Therefore, it was necessary to set a maximum acceptable delay to detection compared with the portfolio that offered the shortest delay to detection. Once a constraint of that nature was added to the decision rule, net benefit and benefit/cost ratio calculations produced very similar portfolios, which were economically much superior to the fastest detection option, while losing very little time until detection occurred.

The study also confirmed that optimal surveillance portfolios indeed outperformed single component surveillance approaches. For instance, the surveillance portfolio with the fastest detection would take 96 days to detect the disease while the fastest single component surveillance strategy would take 184 days

on average; and the portfolio with the highest net benefit would save $188,516 compared with the most economically efficient single component surveillance option (Table 8-13 on Page 239).

The study enriches the understanding of what constitutes the "most optimal" surveillance program. It shows that decision making for surveillance needs to be a multi-criteria process, with different considerations being weighed up according to the nature of the disease and factors such as the seriousness of the clinical effects in people. Each criterion has its practical implications. Sometimes, for certain diseases such as Ebola Hemorrhagic Fever, design of a surveillance program should aim at achieving the highest technical efficiency in detecting the disease. This means the disease needs to be detected as early as possible because of the speed at which it spreads and the severe social and economic consequences it can generate. However, in a less rapidly spreading disease, such as Austeria, by accepting a five day average delay (5.5%) in disease detection compared with that of the portfolio with the fastest detection time, the portfolio with highest net benefit that meets the delay requirement would save $555, 218 (11.5%) compared with the fastest option. Policy makers may well be willing to accept this option balancing different objectives in most cases.

In summary, the study introduced a practical way for screening and identifying most suitable surveillance options under different situations. It demonstrated that the best multi-component surveillance portfolios could perform better than the best single surveillance component, whether evaluated technically or economically. It also indicated the importance of understanding the epidemiology of a disease in identification of the most efficient disease surveillance and mitigation strategies. The approach employed and the OptiSurv tool used in the study are generalizable for optimization of surveillance programs for other zoonotic diseases.

9.8. Practical value of HandiResponse

HandiResponse has the following merits: (i) all the critical steps for designing an optimal risk-based surveillance portfolio are included in the package. With the help of the program, not only could surveillance programs be designed, but also their effects could be assessed ex-ante. (ii) the program is designed to present the inputs, process and the results of each step in a transparent and comprehensible manner, and allows participation from stakeholders in the critical steps such as risk factor identification, estimation of economic consequences of a disease, decision making on most suitable surveillance portfolios; (iii) to address the issue of data scarcity, two approaches have been proposed. The program has compiled and will update periodically remote sensed data sets relevant to disease risk prediction. These data sets will offer a valuable complement to national statistical data sets and *ad hoc* field survey data. They have been standardized and are obtained from source organizations which ensure that the data is of good quality. Moreover, they can be used for any user-defined geographic areas without worrying about national borders. Further, using the program can help countries identify gaps in critical

information needed for surveillance program design. Expert opinions can also be used to guide decisions through the selection of variables and interpretation of available information; (iv) the program is currently still under development with respect to achieving completely user-friendly design, but now that the basic principles have been established, user experience can be enhanced as resources become available. The program operates within the Incident Response Information System (IRIS) that is also a platform for data management in relation to disease investigation and management. A Wiki has been developed for easy reference and consultation, plus, on-line advice on how to use the program is available and (v) HandiResponse program is free of charge.

Last but not least, the program has been specifically designed to handle infectious diseases affects both human and animal populations, hence zoonotic diseases can be mapped, modelled and their surveillance programs designed and assessed. OptiSurv can help identify both technically and economically efficient surveillance portfolios. These portfolios satisfy different policy targets.

9.9. Limitations and future direction

Within this project, HandiResponse has been applied to two well-known diseases (directly transmitted HPAIV H5N1 and vector-borne CCHF) to test its robustness and scope of coverage of disease epitypes. In the future, the program needs to be challenged and validated with additional well-documented infectious disease outbreaks or epidemics. All the other epitypes and sub-epitypes need to be fully covered in the software, and a procedure needs to be implemented in the software to help the user select which epitype template to use for the task they are undertaking. The software also needs an interface which assists user interaction, and more fully integrates the sequence of steps needed to design and implement a surveillance portfolio. Consideration also needs to be given as to how to handle diseases which vary considerably between seasons or have other forms of temporal variation that would influence surveillance decisions

A number of enhancements will be needed to further improve user-friendliness of the existing modules so that a fully user-friendly version of HandiResponse is readily available. Firstly, the modules that currently are in MS-Excel format need to be converted to a web-based version and all modules should have consistent user interfaces. Secondly, transfer of information between successive modules should be automated as far as possible, and should include error detection and reporting, to deal with cases where a user has formulated data in a way which is incompatible with the next module. Thirdly, additional demonstration risk maps need to be developed as templates under each epitype to assist users to make full use of the capabilities of HandiMap. Fourthly, a module on disease mitigation needs to be developed since surveillance needs to be linked to subsequent mitigation, and they must be considered and assessed simultaneously. However this is a major task comparable in size to this project. Lastly, the HandiSurv module needs further refinement so that it can help guide the user through selection of

surveillance approaches by illustrating details on the structure of potential surveillance programs: by whom (institution and personnel), when (frequency and duration) and what (planning, training and supervision, and with what means of implementation).

9.10. Conclusion

A desirable disease surveillance program needs to be technically sound and economically efficient. A surveillance portfolio which makes use of both risk-based and classical surveillance methods represents a promising solution. HandiResponse is a practical tool that could promote the implementation of risk-based surveillance approaches, and could improve both technical and economic efficiency of surveillance programs for infectious diseases, in particular those which affect both human and animals. The project reported in this thesis has demonstrated the practical value of HandiResponse in completing the various stages in the process of designing and implementing a surveillance portfolio.

9.11. References

Alba, A., Casal, J., Napp, S., & Martin, P. (2010). Assessment of different surveillance systems for avian influenza in commercial poultry in Catalonia (North-Eastern Spain). *Preventive Veterinary Medicine, 97*(2), 107-118.

Atkinson, B., Chamberlain, J., Jameson, L. J., Logue, C. H., Lewis, J., Belobrova, E. A., . . . Hewson, R. (2013). Identification and analysis of Crimean-Congo hemorrhagic fever virus from human sera in Tajikistan. *International Journal of Infectious Diseases, 17*(11), e1031-e1037.

Barnabas, R. V., Laukkanen, P., Koskela, P., Kontula, O., Lehtinen, M., & Garnett, G. P. (2006). Epidemiology of HPV 16 and cervical cancer in Finland and the potential impact of vaccination: mathematical modelling analyses. *PLoS Med, 3*(5), e138.

Bettencourt, L. M., & Ribeiro, R. M. (2008). Real time bayesian estimation of the epidemic potential of emerging infectious diseases. *PLoS One, 3*(5), e2185.

Brahmbhatt, M. (2006). *Economic impacts of avian influenza propagation.* Paper presented at the First International Conference on Avian Influenza in Humans.

Brisson, M., Edmunds, W., Gay, N., Law, B., & De Serres, G. (2000). Modelling the impact of immunization on the epidemiology of varicella zoster virus. *Epidemiology and infection, 125*(03), 651-669.

Brownstein, J. S., Freifeld, C. C., & Madoff, L. C. (2009). Digital disease detection—harnessing the Web for public health surveillance. *New England Journal of Medicine, 360*(21), 2153-2157.

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. doi: 10.1016/j.prevetmed.2012.01.009

Cao, C., Xu, M., Chang, C., Xue, Y., Zhong, S., Fang, L., . . . He, Q. (2010). Risk analysis for the highly pathogenic avian influenza in Mainland China using meta-modelling. *Chinese Science Bulletin, 55*(36), 4168-4178.

Carroll, L. N., Au, A. P., Detwiler, L. T., Fu, T.-c., Painter, I. S., & Abernethy, N. F. (2014). Visualization and analytics tools for infectious disease epidemiology: A systematic review. *Journal of Biomedical Informatics, 51*, 287-298.

Christensen, J., El Allaki, F., & Vallières, A. (2014). Adapting a scenario tree model for freedom from disease as surveillance progresses: the Canadian notifiable avian influenza model. *Preventive veterinary medicine, 114*(2), 132-144.

Eidson, M., Kramer, L., Stone, W., Hagiwara, Y., Schmit, K., & Team, N. Y. S. W. N. V. A. S. (2001). Dead bird surveillance as an early warning system for West Nile virus. *Emerging infectious diseases, 7*(4), 631.

Fang, L.-Q., de Vlas, S. J., Liang, S., Looman, C. W., Gong, P., Xu, B., . . . Cao, W.-C. (2008). Environmental factors contributing to the spread of H5N1 avian influenza in mainland China. *PLoS One, 3*(5), e2268.

Gilbert, M., Xiao, X., Pfeiffer, D. U., Epprecht, M., Boles, S., Czarnecki, C., . . . Otte, M. J. (2008). Mapping H5N1 highly pathogenic avian influenza risk in Southeast Asia. *Proceedings of the National Academy of Sciences, 105*(12), 4769-4774.

Guo, X., Claassen, G. D. H., Oude Lansink, A. G. J. M., & Saatkamp, H. W. (2014). A conceptual framework for economic optimization of single hazard surveillance in livestock production chains. *Preventive veterinary medicine, 114*(3-4), 188-200. doi: 10.1016/j.prevetmed.2014.02.003

Hadorn, D. C., Rufenacht, J., Hauser, R., & Stark, 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. doi: 10.1016/s0167-5877(02)00193-9

Hadorn, D. C., & Stark, K. D. C. (2008). Evaluation and optimization of surveillance systems for rare and emerging infectious diseases. *Veterinary research, 39*(6). doi: 10.1051/vetres:2008033

Hay, S. I., Battle, K. E., Pigott, D. M., Smith, D. L., Moyes, C. L., Bhatt, S., . . . George, D. B. (2013). Global mapping of infectious disease. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 368*(1614), 20120250.

Hewson, R. (2013). Risk analysis of Crimean Congo Haemorrhagic Fever in Mongolia & Surveillance planning: Public Health England - Microbiology Services.

Heymann, D. L., & Rodier, G. R. (1998). Global surveillance of communicable diseases. *Emerging infectious diseases, 4*(3), 362.

Hoogstraal, H. (1979). Review Article: The Epidemiology of Tick-Borne Crimean-Congo Hemorrhagic Fever in Asia, Europe, and Africa. *Journal of medical entomology, 15*(4), 307-417.

Kahn, L. H. (2006). Confronting zoonoses, linking human and veterinary medicine. *Emerging infectious diseases, 12*(4), 556.

Keeling, M. J., Woolhouse, M. E., Shaw, D. J., Matthews, L., Chase-Topping, M., Haydon, D. T., . . . Grenfell, B. T. (2001). Dynamics of the 2001 UK foot and mouth epidemic: stochastic dispersal in a heterogeneous landscape. *Science, 294*(5543), 813-817.

Kitron, U. (2000). Risk maps: transmission and burden of vector-borne diseases. *Parasitol Today, 16*(8), 324-325.

Koopman, J. S. (2005). Infection transmission science and models. *Japanese journal of infectious diseases, 58*(6), S.

Kuiken, T., Leighton, F., Fouchier, R., & LeDuc, J. (2005). Pathogen surveillance in animals. *Science, 309*(5741), 1680.

Kulasekera, V. L., Kramer, L., Nasci, R. S., Mostashari, F., Cherry, B., Trock, S. C., . . . Miller, J. R. (2001). West Nile virus infection in mosquitoes, birds, horses, and humans, Staten Island, New York, 2000. *Emerging infectious diseases, 7*(4), 722.

Kulldorff, M., Fang, Z., & Walsh, S. J. (2003). A tree-based scan statistic for database disease surveillance. *Biometrics, 59*(2), 323-331. doi: 10.1111/1541-0420.00039

Lawson, A. B., & Leimich, P. (2000). Approaches to the space-time modelling of infectious disease behaviour. *Mathematical Medicine and Biology, 17*(1), 1-13.

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

Martin, V., Pfeiffer, D. U., Zhou, X., Xiao, X., Prosser, D. J., Guo, F., & Gilbert, M. (2011). Spatial distribution and risk factors of highly pathogenic avian influenza (HPAI) H5N1 in China. *PLoS Pathog, 7*(3), e1001308.

Merler, S., Ajelli, M., Fumanelli, L., Gomes, M. F., y Piontti, A. P., Rossi, L., . . . Vespignani, A. (2015). Spatiotemporal spread of the 2014 outbreak of Ebola virus disease in Liberia and the effectiveness of non-pharmaceutical interventions: a computational modelling analysis. *The Lancet Infectious Diseases, 15*(2), 204-211.

Morse, S. S. (2007). Global infectious disease surveillance and health intelligence. *Health Affairs, 26*(4), 1069-1077.

Mossong, J., Hens, N., Jit, M., Beutels, P., Auranen, K., Mikolajczyk, R., . . . Wallinga, J. (2008). Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med, 5*(3), e74.

Mostashari, F., Bunning, M. L., Kitsutani, P. T., Singer, D. A., Nash, D., Cooper, M. J., . . . Fine, A. D. (2001). Epidemic West Nile encephalitis, New York, 1999: results of a household-based seroepidemiological survey. *The Lancet, 358*(9278), 261-264.

Mostashari, F., Kulldorff, M., Hartman, J. J., Miller, J. R., & Kulasekera, V. (2003). Dead bird clusters as an early warning system for West Nile virus activity. *Emerging infectious diseases, 9*(6), 641.

Myers, M. F., Rogers, D., Cox, J., Flahault, A., & Hay, S. (2000). Forecasting disease risk for increased epidemic preparedness in public health. *Advances in Parasitology, 47*, 309-330.

Narrod, C., Zinsstag, J., & Tiongco, M. (2012). A one health framework for estimating the economic costs of zoonotic diseases on society. *EcoHealth, 9*(2), 150-162.

Ortiz-Pelaez, A., Pfeiffer, D. U., Soares-Magalhaes, R. J., & Guitian, F. J. (2006). Use of social network analysis to characterize the pattern of animal movements in the initial phases of the 2001 foot and mouth disease (FMD) epidemic in the UK. *Preventive veterinary medicine, 76*(1-2), 40-55. doi: 10.1016/j.prevetmed.2006.04.007

Paul, M., Held, L., & Toschke, A. M. (2008). Multivariate modelling of infectious disease surveillance data. *Statistics in Medicine, 27*(29), 6250-6267. doi: 10.1002/sim.3440

Peterson, A. T., & Williams, R. A. (2008). Risk mapping of highly pathogenic avian influenza distribution and spread. *Ecology and Society, 13*, 15.

Pfeiffer, D. U., Minh, P. Q., Martin, V., Epprecht, M., & Otte, M. J. (2007). An analysis of the spatial and temporal patterns of highly pathogenic avian influenza occurrence in Vietnam using national surveillance data. *The Veterinary Journal, 174*(2), 302-309.

Polgreen, P. M., Chen, Z., Segre, A. M., Harris, M. L., Pentella, M. A., & Rushton, G. (2009a). Optimizing influenza sentinel surveillance at the state level. *American journal of epidemiology*, 1300-1306.

Prattley, D. J., Morris, R. S., Stevenson, M. A., & Thornton, R. (2007). Application of portfolio theory to risk-based allocation of surveillance resources in animal populations. *Preventive veterinary medicine, 81*(1-3), 56-69. doi: 10.1016/j.prevetmed.2007.04.009

Presi, P., Staerk, K. D. C., Knopf, L., Breidenbach, E., Sanaa, M., Frey, J., & Regula, G. (2008). Efficiency of risk-based vs. random sampling for the monitoring of tetracycline residues in slaughtered calves in Switzerland. *Food Additives and Contaminants, 25*(5), 566-573. doi: 10.1080/02652030701660544

Reist, M., Jemmi, T., & Staerk, K. D. C. (2012). Policy-driven development of cost-effective, risk-based surveillance strategies. *Preventive veterinary medicine, 105*(3), 176-184. doi: 10.1016/j.prevetmed.2011.12.014

Riley, S. (2007). Large-scale spatial-transmission models of infectious disease. *Science, 316*(5829), 1298-1301.

Rivas, A. L., Chowell, G., Schwager, S., Fasina, F. O., Hoogesteijn, A. L., Smith, S. D., . . . Hyman, J. M. (2010). Lessons from Nigeria: the role of roads in the geo-temporal progression of avian influenza (H5N1) virus. *Epidemiology and infection, 138*(02), 192-198.

Scarpino, S. V., Dimitrov, N. B., & Meyers, L. A. (2012). Optimizing provider recruitment for influenza surveillance networks. *PLoS Comput Biol, 8*(4), e1002472.

Schwermer, H., Reding, I., & 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.

Sparks, R. (2013). Challenges in designing a disease surveillance plan: What we have and what we need? *IIE Transactions on Healthcare Systems Engineering, 3*(3), 181-192.

Stark, K. D. C., Regula, G., Hernandez, J., Knopf, L., Fuchs, K., Morris, R. S., & 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), (28 February 2006)-(2028 February 2006).

Tatem, A. J., Smith, D. L., Gething, P. W., Kabaria, C. W., Snow, R. W., & Hay, S. I. (2010). Ranking of elimination feasibility between malaria-endemic countries. *The Lancet, 376*(9752), 1579-1591.

Tavornpanich, S., Gardner, I. A., Carpenter, T. E., Johnson, W. O., & Anderson, R. J. (2006). Evaluation of cost-effectiveness of targeted sampling methods for detection of Mycobacterium aviumsubsp paratuberculosis infection in dairy herds. *American journal of veterinary research, 67*(5), 821-828.

Tiensin, T., Chaitaweesup, P., Songserm, T., Chaising, A., Hoonsuwan, W., Buranathai, C., . . . Gilbert, M. (2005). Highly pathogenic avian influenza H5N1. *Emerging infectious diseases, 11*, 1664-1672.

Tsai, P., Scott, K. A., Pappaioanou, M., Gonzalez, M. C., & Keusch, G. T. (2009). *Sustaining global surveillance and response to emerging zoonotic diseases*: National Academies Press.

UNDG. (2015). Socio-Economic Impact of Ebola Virus Disease in West African Countries - A call for national and regional containment, recovery and prevention (pp. 93). New York, USA: United Nations Development Programme Regional Bureau for Africa.

Walsh, D. P., & Miller, M. W. (2010). A Weighted Surveillance Approach for Detecting Chronic Wasting Disease Foci. *Journal of wildlife diseases, 46*(1), 118-135.

Wieland, S. C., Brownstein, J. S., Berger, B., & Mandl, K. D. (2007). Automated real time constant-specificity surveillance for disease outbreaks. *Bmc Medical Informatics and Decision Making, 7*. doi: 10.1186/1472-6947-7-15

Willeberg, P., Nielsen, L. R., & 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. doi: 10.1016/j.prevetmed.2011.12.012

Williams, R. A., & Peterson, A. T. (2009). Ecology and geography of avian influenza (HPAI H5N1) transmission in the Middle East and northeastern Africa. *International Journal of Health Geographics, 8*(1), 47.

Woolhouse, M. (2011). How to make predictions about future infectious disease risks. *Philosophical Transactions of the Royal Society B: Biological Sciences, 366*(1573), 2045-2054.

World Bank. (2014). The economic impact of the 2014 Ebola epidemic : short and medium term estimates for West Africa. . In W. B. Group (Ed.), (pp. 71). Washington, DC.

Yashina, L., Petrova, I., Seregin, S., Vyshemirskii, O., Lvov, D., Aristova, V., . . . Kuzina, I. (2003). Genetic variability of Crimean-Congo haemorrhagic fever virus in Russia and Central Asia. *Journal of General Virology, 84*(5), 1199-1206.

Yen, Y., Kong, L., Lee, L., Zhang, Y., Li, F., Cai, B., & Gao, S. (1985). Characteristics of Crimean-Congo hemorrhagic fever virus (Xinjiang strain) in China. *The American journal of tropical medicine and hygiene, 34*(6), 1179-1182.

Yupiana, Y., de Vlas, S. J., Adnan, N. M., & Richardus, J. H. (2010). Risk factors of poultry outbreaks and human cases of H5N1 avian influenza virus infection in West Java Province, Indonesia. *International Journal of Infectious Diseases, 14*(9), e800-e805.

Zhang, L., Jamal, I., Qu, Y., & Zeng, G. (2011). Optimization of event detection methods for disease surveillance. *Epidemiologie et Sante Animale*(59/60), 196-198.