QUANTITATIVE EVALUATION OF THE EFFECTIVENESS OF OFFICIAL ANIMAL DISEASE SURVEILLANCE PROGRAMMES

Sarah Welby

GHENT UNIVERSITY

FACULTY OF VETERINARY MEDICINE
accredited by EAEVE

CODA - CERVA

QUANTITATIVE EVALUATION OF THE EFFECTIVENESS OF OFFICIAL ANIMAL DISEASE SURVEILLANCE PROGRAMMES

Dissertation submitted in fulfilment of the requirements for the degree of Doctor (PhD) in Veterinary Sciences

Sarah Welby

Epidemiology, risk assessemnt and surveillance unit

Veterinary Agrochemical Research Centre (CODA-CERVA)

Belgium

Promotors:

Prof Dr Jeroen Dewulf, Faculteit Diergeneeskunde, UGent

Dr Yves Van der Stede, European Food Safety Agency, EFSA, Parma

Exam committee:

Prof Dr Richard Ducatelle, Faculteit Diergeneeskunde, UGent

Prof Dr Pascal Hendrikx, French Agency for Food, Environmental and Occupational Health & Safety, ANSES, France

Prof Dr Sarah Gabriël, Faculteit Diergeneeskunde, UGent

Dr Jozef Hooyberghs, Federal Agency for the Safety of the Food Chain, FAFSC, Brussels

Dr Koen Mintiens, Boerenbond, Leuven

Dr Steven Sarrazin, Faculteit Diergeneeskunde, Ugent

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

A

Ab Abortion surveillance component

AR Adjusted risk of infection

ASe Sensitivity at animal level

AI Avian influenza

B

BB Bovine brucellosis

BE Belgium

BTB Bovine tuberculosis

BTV Bluetongue virus

C

CSe Surveillance component sensitivity

D

Da Dairy surveillance component

E

EPI_a EPI_h Effective probability of infection at animal (a) or herd level (h)

EPD_a EPD_h Effective probability of detection at animal (a) or herd level (h)

F

FC Fattening calves

FR France

G

GEE Generalised estimating equation model

H

HSe Sensitivity at herd level

I

ID Intradermal tuberculin skin test

IF Import from officially free countries surveillance component

IMP Import surveillance component

IR Import from non-officially free countries surveillance component

N

n sample population size

N representative population size

NL The Netherlands

P

P* P_a P_h Design prevalence (P*) at animal (a) or herd level (h)

PIntro Probability of introduction

PM Post mortem visual inspection

PostPFree Posterior confidence of population freedom (= negative predictive value)

PostPInf Posterior probability of being infected = 1 – posterior confidence of freedom (NPV)

PPr Representative population proportion

PriorPFree Prior confidence of population freedom before undertaking current surveillance

PriorPInf Prior probability of being infected = 1 – prior confidence of freedom

PUR Purchase surveillance component

R

RR Relative risk

S

SLGH Slaughterhouse surveillance component

SSC Surveillance system component

SSe Surveillance system sensitivity

SPr Sampled population proportion

T

Tse Diagnostic test sensitivity

Tr Trade surveillance component

W

WS Winter screening surveillance component

SUMMARY

National and international standards have ruled notifiable animal disease surveillance programmes for the last 100 years. Over the past two decades a shift from input based standards towards output based standards is observed with an obligation of results (i.e. achieve given objectives) rather than means (fixed sample size, diagnostic tests). This major shift towards the implementation of a more 'fit for purpose' surveillance programme design enabled taking into account the specific disease as well as the country specific characteristics, risk factors and husbandry practices. Because animal health surveillance reliability is a major concern for free trade as well as for consumer trust, proving effectiveness and efficiency of surveillance remains a constant challenge. The current thesis aims at describing and applying scenario tree methods (output based methods) for different surveillance programmes in Belgium as a tool to assess effectiveness and efficiency of these surveillance systems. The overarching objective across all studies in the current thesis was to provide an extensive overview of the possibilities offered by these methods as well as discuss their limitations and assumptions. This thesis aims to provide a conceptual framework for assessing effectiveness and efficiency of notifiable animal disease surveillance programmes.

A first study aimed at redesigning the Belgian avian influenza active surveillance programme for domestic birds in professional poultry holdings based on a risk analysis approach. Stochastic quantitative analysis was run to obtain effective probabilities of infection and sensitivity estimates for the detection of an infected bird in the different risk groups identified, using national animal identification registries, together with past outbreak data found in literature. An optimal number of holdings for each risk group was then estimated on the basis of the different effective probabilities of infection and sensitivities obtained. In certain risk groups the optimized sampling size decreased drastically such as in chicken farms without outdoor facilities in non-risk zones while in other risk groups the initial sample size was underestimated such as for chickens raised in risk zones with outdoor facilities. In addition it was recommended to focus surveillance in certain seasons as well as increase frequency of sampling in certain risk groups in order to allow early detection.

This study was a useful tool for decision makers to reallocate the total amount of samples to be taken in the next year(s) in Belgium, thus optimizing the field resources, and improving efficiency of disease surveillance such as required by international standards.

The second study investigated different surveillance components of the bovine brucellosis surveillance system in Belgium and how this surveillance programme could be optimized to ensure freedom of infection in Belgium. The current surveillance system (prior to 2009) and the impact of reducing surveillance in certain risk groups on the whole surveillance system sensitivity were simulated together stochastically. Results demonstrated that a significant decrease in the total number of samples was feasible whilst maintaining a 99% confidence level of disease freedom. The ideal number of samples was estimated to be around 30,000 (thus a 30 fold decrease from pre 2009 levels) under the condition that the new surveillance would target risk groups. Hence increasing the number of reported and tested abortions by farmers and veterinarians was necessary to reach a minimal annual number of reports of 8,000 abortions instead of 4,000. The reporting of abortion was strongly stimulated, via the abortion protocol. Furthermore, it was advised to test a random selection of 15,000 cattle from all herds which did not declare any abortions during the past years, these cattle were investigated for bovine brucellosis during winter screening. In addition, it was recommended to test 8,500 cattle during purchase (randomly selected in regional laboratories) and 8,500 cattle from intra community trade from officially free countries, as well as all cattle from intra community trade from non-officially free countries. The new surveillance programme led to the detection of one case in 2010 following an abortion investigation.

The third study used scenario tree modelling to evaluate and to identify alternatives to optimize the bovine tuberculosis surveillance system in accordance with European legislation (Council Directive 64/432/EEC). Data from 2005-2009 regarding cattle population, movement and surveillance were collected to feed the stochastic scenario tree simulation model, using 10,000 iterations per simulation. A total of 7,403,826 cattle movement history records were obtained for the 2,678,020 cattle from 36,059 cattle herds still active in 2009. The sensitivity analysis showed that the most influential input parameter explaining the variability around the output came

from the uncertainty distribution around the sensitivity of the diagnostic tests used within the bTB surveillance. Provided all animals are inspected and post mortem inspection is highly sensitive, slaughterhouse surveillance was the most effective surveillance component. However it also highlighted that in 25% of the cases an infection could go unnoticed using only slaughterhouse surveillance and it was therefore advised to increase the sensitivity of surveillance in other components as well. Using for instance different diagnostic tests instead of the intradermal skin test during winter screening, purchases and import would increase greatly the sensitivity and the confidence level of Belgium's posterior probability of freedom from bTB infection status (negative predictive value of the surveillance system).

The aim of the fourth study was to assess the sensitivity of the four major bluetongue surveillance components implemented in Belgium in 2007 for farmed animals and prescribed by the European Union regulation: winter serological screening, sentinel system, passive clinical surveillance and export testing. Scenario tree methodology was used to evaluate the relative sensitivity of detection and targeted approach of each component in terms of early detection and freedom of infection substantiation. Field data collected from the previous years' outbreaks in Belgium were used to determine the risk groups to be considered. The best sensitivities at herd level, taking into account the diagnostic test sensitivity, design prevalence and the number of animals tested within a herd were obtained with the winter screening and sentinel component. At component level, sensitivity analysis showed that this sensitivity would be all the more increased if targeted towards the right risk groups. Despite the very low herd sensitivity, passive clinical surveillance turned out to be the most sensitive component in terms of early detection, under the condition of high disease awareness, absence of vaccination and visible clinical signs. The conclusions of this study showed that passive clinical surveillance could be suitable for early detection and substantiation of freedom but because of its limitations it was all together advised to maintain a repeated cross sectional monthly survey or 4 months period during the vector season and in certain risk groups for substantiation of disease freedom.

The last study compared the relative efficacy and cost efficiency of different surveillance components in proving the absence of infection (freedom of infection) or

early detection of vector-borne diseases in cattle populations within different countries (using bluetongue as case study). Bluetongue outbreak and surveillance data from the epidemic that occurred in the Netherlands, France and Belgium, in the years 2006 and 2007 provided reliable input data for evaluating the different surveillance components in place in each country. In a first stage, the different surveillance components (sentinel, yearly cross sectional and passive clinical reporting) in 2006-2007 within each country were evaluated in terms of efficacy for substantiating freedom of infection. The yearly cross sectional survey and passive clinical reporting within each country performed well with sensitivity of detection values ranging around 0.99. The large number of cattle and herds sampled and the high diagnostic tests sensitivity contributed to these high sensitivity values within each of the surveillance components. In a second stage, it was investigated how effective the components were for (early)-detection and whether syndromic surveillance using reproductive performance data, milk production and mortality data could be of an added value. To account for the timeliness of detection, epidemic curves obtained from Reed-Frost models were used to estimate the input to feed the scenario tree models. Depending on the expected within herd prevalence, passive clinical reporting and syndromic surveillance performed much better than the active components, with a 0.99 median sensitivity value. Depending on the assumed direct costs of running each of these surveillance components, passive clinical surveillance together with syndromic surveillance (based on reproductive performance data) turned out to be the most cost efficient for detection of vector-borne diseases such as bluetongue. To conclude, for emerging or re-emerging vector-borne disease such as bluetongue, it is recommended to use passive clinical and syndromic surveillance as early detection systems for maximum cost efficiency. Once an infection is detected, a cross sectional screening is recommended for substantiating freedom of infection. Sentinel surveillance is useful for monitoring the disease evolution provided sufficient herds are sampled and risk based targeted.

When evaluating surveillance programs, the key questions are: (i) how better surveillance could be achieved with less money; and (ii) is surveillance effective as well as worth its cost?

To conclude tools explored during this thesis revealed interesting perspectives for evaluation and design of surveillance taking into account Belgian country specific data and risk factors. Alternatives explored enable conciliating different objectives encountered by the different animal heath surveillance industry and stakeholders, thereby increasing compliance and acceptability of surveillance. Cost could be optimised by improving sensitivity while taking into account specificity not only to improve confidence level in freedom but enhancing disease detection too. In addition, consumers and various partners in animal health become ever more concerned with the reliability of surveillance effectiveness and its cost efficiency. Local initiatives to improve surveillance of non-notifiable diseases, for which mutual trust is desirable to allow free trade but as well to implement sustainable production, safe public health and economy. The tools described in the current thesis do certainly offer promising perspectives to assess cost efficiency of other non-regulated diseases and sustainable one health surveillance across many levels and sectors encompassed by surveillance.

SAMENVATTING

Nationale en internationale normen hebben de afgelopen 100 jaar de inhoud bepaald van officiële dierziektenprogramma's. In de laatste twee decennia vond een shift in de wetgeving plaats van een verplichting tot het behalen van resultaten in plaats van middelen (vaste steekproefgrootte, diagnostische testen). Deze belangrijke verandering naar een meer 'fit for purpose' surveillance design maakte het mogelijk rekening te houden met de ziekte- en landspecifieke kenmerken, risicofactoren en veehouderijpraktijken. Omdat de betrouwbaarheid van het diergezondheidsstatuut 'van belang is voor vrijhandel en consumentenvertrouwen, blijft de effectiviteit en de doeltreffendheid van de surveillance een constante uitdaging. Het huidige proefschrift heeft tot doel om scenario-tree methodologie (output-gebaseerde methoden) te beschrijven en te gebruiken voor verschillende aangifteplichtige dierziekten en toe te lichten hoe deze instrumenten kunnen helpen bij het beoordelen van de effectiviteit en efficiëntie van surveillancesystemen. De overkoepelende doelstelling van dit proefschrift is het geven van een uitgebreid overzicht van de mogelijkheden die deze methoden bieden en om hun beperkingen en aannames te bespreken. Hiermee beoogt dit proefschrift een conceptueel model te bieden voor de beoordeling van de effectiviteit en efficiëntie van bewakingsprogramma's voor de officiële aangifteplichtige dierziekten.

De eerste studie was gericht op het herontwerpen van het Belgische actieve avian influenza bewakingsprogramma in professionele pluimveebedrijven, op basis van een risicoanalyse-aanpak. Een kwantitatieve stochastische analyse werd uitgevoerd om effectieve waarschijnlijkheden van infectie en schattingen van de gevoeligheid voor detectie van een besmette vogel in de verschillende geïdentificeerde risicogroepen te bekomen, met behulp van nationale registers voor dierenidentificatie en gepubliceerde gegevens van vroegere uitbraken. Een optimaal aantal te bemonsteren bedrijven werdgeraamd voor elke risicogroep op basis van de verkregen effectieve waarschijnlijkheden van infectie en gevoeligheden. In bepaalde risicogroepen daalde de geoptimaliseerde steekproefgrootte drastisch t.o.v. de oorspronkelijke steekproefgrootte zoals bij kippenboerderijen zonder openluchtfaciliteiten in niet-risicogebieden, terwijl in andere risicogroepen de

aanvankelijke steekproefgrootte onderschat werd zoals bij kippen in risicogebieden met openluchtfaciliteiten. Daarnaast werd aanbevolen om in bepaalde seizoenen het toezicht aan te scherpen en de frequentie van steekproeven in bepaalde risicogroepen te verhogen om vroegtijdige detectie mogelijk te maken. Deze studie was een nuttig instrument voor beleidsmakers om de totale hoeveelheid van de, over het (de) komende ja(a)r(en) te nemen monsters in België ,te herverdelen , en zodoende de middelen op het terrein te optimaliseren en de efficiëntie van ziektebewaking te verbeteren, zoals vereist door de internationale normen.

In het tweede onderzoek werden verschillende bewakings componenten van het boviene brucellosis surveillancesysteem in België onderzocht en werd bekeken hoe dit surveillanceprogramma optimaal geoptimaliseerd zou kunnen worden om de vrijheid van infectie in België te waarborgen. Het huidige bewakingssysteem (vóór 2009) en de impact van het verminderen van bewaking in bepaalde risicogroepen op de gevoeligheid van het gehele bewakingssysteem werden samen stochastisch gesimuleerd. Resultaten toonden aan dat een significante daling van het totale aantal monsters haalbaar was en toch een 99% betrouwbaarheidsniveau van ziektevrijheid behouden bleef. Het ideale aantal monsters werd geschat op ongeveer 30.000 (dus een 30-voudige afname van wat voor 2009 gedaan werd) onder de voorwaarde dat het nieuwe toezicht zich op risicogroepen zou richten. Zo was het verhogen van het aantal gerapporteerde en geteste abortussen door de boeren en dierenartsen noodzakelijk om een minimaal jaarlijks aantal meldingen van 8.000 abortussen te bereiken in plaats van 4.000. De melding van abortus werd sterk gestimuleerd via het abortusprotocol. Verder werd geadviseerd om een willekeurige selectie van 15.000 runderen te testen uit alle kuddes die de afgelopen jaren geen abortussens hebben aangetoond. Daarnaast werd aanbevolen om 8.500 runderen te testen tijdens de aankoop (willekeurig geselecteerd in regionale laboratoria) en 8.500 runderen uit de intracommunautaire handel van officieel-vrije landen, evenals alle runderen uit de intracommunautaire handel uit niet-officieel vrije landen. Het nieuwe bewakingsprogramma leidde tot de opsporing van één geval in 2010 na een abortusonderzoek.

De derde studie gebruikte scenario-modellering voor de evaluatie van de verschillende componenten van het huidige boviene tuberculose

surveillanceprogramma, en voor de identificatie van alternatieven voor optimalisatie van dit programma, in overeenstemming met de Europese wetgeving (Europese richtlijn 64/432/EEC). Gegevens van 2005-2009 betreft de runderpopulatie, beweging en bewaking werden verzameld om het stochastische scenario boom-simulatiemodel te voeden, met behulp van 10.000 iteraties per simulatie. In totaal werden 7.403.826 historische dierbewegingsrecords verkregen voor de 2.678.020 runderen uit 36.059 veestapels die in 2009 actief waren. De gevoeligheidsanalyse toonde aan dat de meest invloedrijke invoerparameter die de variabiliteit rond de output verklaart, afkomstig was van de onzekerheidsverdeling rond de gevoeligheid van de diagnostische tests die worden gebruikt binnen de bTB-surveillance. Op voorwaarde dat alle dieren onderzocht werden en het post-mortem onderzoek zeer gevoelig is, was de slachthuisbewaking de meest effectieve bewakingscomponent. Als niet aan deze voorwaarden werd voldaan, was de onzekerheid over de gemiddelde gevoeligheid van deze component belangrijk. Toch bleek ook dat in 25% van de gevallen een infectie onopgemerkt zou kunnen blijven met enkel slachthuisbewaking; daarom werd aangeraden de gevoeligheid van toezicht in andere componenten te vergroten. Met behulp van bijvoorbeeld verschillende diagnostische tests in plaats van de intradermale huidtest, onderzoek bij aankoop en invoer, kan de gevoeligheid en het vertrouwensniveau sterk verhogen (negatief voorspellende waarde van het surveillancesysteem).

De doelstelling van het vierde onderzoek was het beoordelen van de gevoeligheid van de vier grote blauwtong surveillancecomponenten die in België in 2007 voor dieren werden geïmplementeerd, zoals voorgeschreven door de EU-verordening: winter serologische screening, sentinel systeem, passief klinisch toezicht en export testen. De scenario boommethode werd gebruikt om de relatieve gevoeligheid van detectie en de gerichte benadering van elke component te evalueren in termen van vroegtijdige detectie en vrijheid van infectieonderbouw. Veldgegevens verzameld in België tijdens de uitbraken van de vorige jaren werden gebruikt om de risicogroepen te bepalen die overwogen zouden moeten worden. De beste gevoeligheid op beslagniveau, rekening houdend met de diagnostische testgevoeligheid, de ontwerpprevalentie en het aantal dieren dat in een beslag werd getest, werd verkregen bij de winterscreening en de sentinel component. Op component niveau,

toonde de gevoeligheidsanalyse aan dat deze gevoeligheid nog hoger kan, indien gericht op de juiste risicogroepen. Ondanks de zeer lage beslag niveau gevoeligheid bleek passief klinisch toezicht de meest gevoelige component in termen van vroege detectie, onder de voorwaarde van een hoog ziektebewustzijn, afwezigheid van vaccinatie en zichtbare klinische symptomen. Uit de conclusies van deze studie bleek dat passief klinisch toezicht geschikt zou kunnen zijn voor vroegtijdig opsporen en onderbouwen van vrijheid van ziekte, maar door de beperkingen werd geadviseerd om een herhaalde cross sectionele survey per maand of per 4 maandelijkse periodes gedurende het vectorseizoen en in bepaalde risicogroepen te behouden ter onderbouwing van het vrij zijn van ziekte.

Het laatste onderzoek vergeleek de relatieve doeltreffendheid en kostenefficiëntie van verschillende bewakingscomponenten om de afwezigheid van infectie (infectievrijheid) aan te tonen of om de vroegtijdige detectie van vector-overdraagbare ziekten bij runderpopulaties in verschillende landen toe te staan (met behulp van blauwtong als case study). Blauwtong uitbraak- en bewakingsgegevens van de epidemie die zich in Nederland, Frankrijk en België voordeed in de jaren 2006 en 2007, hebben betrouwbare invoergegevens verstrekt om de verschillende bewakingscomponenten in elk land te evalueren. In een eerste fase werden de verschillende surveillancecomponenten (sentinel, jaarlijkse cross-sectionele studie en passieve klinische rapportage) in 2006-2007 in elk land geëvalueerd in termen van werkzaamheid om de vrijheid van infectie te ondersteunen. De jaarlijkse cross-sectioneele studie en passieve klinische rapportage binnen elk land scoorden goed, met een detectiegevoeligheid rond 0.99. Het groot aantal bemonsterde dieren en beslagen en de hoge gevoeligheid van de diagnostische tests droegen bij aan de hoge gevoeligheidswaarden binnen elk van deze bewakingscomponenten. In een tweede fase werd onderzocht hoe effectief de componenten waren voor (vroege) detectie en of syndroombewaking met behulp van fertiliteitsgegevens, melkproductie- en sterftegegevens van toegevoegde waarde kunnen zijn. Om rekening te houden met de tijdigheid van detectie werden epidemische curves, verkregen uit Reed-Frost-modellen, gebruikt om de input voor de scenario-boommodellen te schatten. Afhankelijk van de verwachte prevalentie binnen de beslagen, bleek passieve klinische rapportage en syndroombewaking veel beter dan de actieve componenten,

met een mediaan van 0.99 voor de gevoeligheidswaarden. Afhankelijk van de veronderstelde directe kosten van het uitvoeren van elk van deze bewakingscomponenten bleek passief klinisch toezicht samen met syndroombewaking (gebaseerd op fertiliteitsgegevens) het meest kostenefficiënt voor het opsporen van vector-overdraagbare ziekten zoals blauwtong. Tenslotte, voor (her)opkomende vector-overdraagbare ziekten zoals blauwtong, is het aanbevolen passieve klinische bewaking en syndroombewaking te gebruiken als vroege detectiesystemen voor een maximale kosten-efficiëntie. Zodra een infectie is gedetecteerd, wordt een cross-sectionele screening aanbevolen om de vrijheid van infectie te ondersteunen. Sentinel surveillance is nuttig om de ziekteontwikkeling te monitoren, op voorwaarde dat voldoende kuddes worden bemonsterd via risico gebaseerde selectie.

Bij het evalueren van bewakingsprogramma's zijn de sleutelvragen: ‘hoe kan betere bewaking worden bereikt met minder geld’ en ‘is het toezicht effectief en het het geld waard’?

Men kan besluiten dat de methoden die tijdens dit proefschrift werden onderzocht interessante perspectieven aan het licht brachtten voor evaluatie en ontwerp van bewaking, rekening houdend met Belgische landspecifieke gegevens en risicofactoren.. De onderzochte alternatieven, maken het mogelijk om verschillende doelstellingen waarmee de verschillende commerciële en andere belanghebbenden in de dierengezondheidsbewaking worden geconfronteerd aan te pakken, waardoor de naleving en acceptatie van het toezicht wordt verhoogd. Kosten kunnen worden geoptimaliseerd door de gevoeligheid te verbeteren, waarbij rekening wordt gehouden met de specificiteit, niet alleen om het betrouwbaarheidsniveau tijdens vrijheid van ziekte te verbeteren maar ook om de detectie van ziekten te verbeteren. Bovendien zijn consumenten en verschillende partners in dierengezondheid steeds meer bezorgd over de betrouwbaarheid van de bewakingsdoeltreffendheid evenals over de kosten-efficiëntie ervan. Lokale initiatieven zijn ontstaan voor het verbeteren van het toezicht op niet-meldingsplichtige ziekten, waarvoor wederzijds vertrouwen wenselijk is om vrijhandel toe te staan, maar ook om duurzame productie te implementeren. De werkwijzen die in het huidige proefschrift zijn beschreven, zouden in deze context inderdaad veelbelovende perspectieven kunnen bieden voor het

beoordelen van de kostenefficiëntie van andere niet-gereguleerde ziekten, en voor een duurzame ‘one health’ gezondheidsbewaking op de verschillende niveaus en in de verschillende sectoren die onder toezicht staan.

CHAPTER I: GENERAL INTRODUCTION

SURVEILLANCE DEFINITION

The word surveillance originates from the French words "*sur*" and "*veiller*" meaning to 'watch over'. The word can take different meanings and sometimes can be mistaken with the word monitoring. However, in the case of animal diseases, surveillance usually signifies the action of collecting information on populations' health status, in order to plan interventions; whereas monitoring means more specifically all activities aimed at detecting changes in the epidemiological parameters of a particular disease, without implementing any control actions (Bisdorff et al., 2016; FAO, 2017; Hoinville et al., 2013). Thus, surveillance involves all activities from the collection and analysis of data to its interpretation, as well as communicating on the nature and implementation of specific intervention strategies whenever a positive signal is detected (Drewe et al., 2012; FAO, 2017, 2011; Hoinville et al., 2013; OIE, 2017a).

Growth in human population together with globalisation, which have resulted in increases of livestock production and animal movements, have both played a major role in the spread of animal diseases. Major epidemics for example, bovine spongiform encephalopathy (BSE), avian influenza (Elbers et al., 2004), foot and mouth (Valarcher et al., 2008), bluetongue (de Koeijer et al., 2006; Méroc et al., 2009), as well as re-emergence of well-known 'old' diseases such as tuberculosis and brucellosis (FAO, 2017, Gilbert et al, 2005), have occurred at the beginning of the 21st century with devastating impact for farmers and economy as a whole. Faced with the new challenges, created by emerging and re-emerging diseases, by increases in demand for food products and by trade across the European Union (EU) or with third world countries, as well as faced with the substantial evolution of science and technology, the need for clear guidelines and rules to guarantee safe feed, food and trade have emerged. Indeed, to allow mutual trust and recognition between the different livestock production stakeholders, and public health partners, evidence of safety is a prerequisite. This can be obtained by mandatory rules established by authorities as well as non-regulatory -local initiatives. In this context, the existing animal health policies drawn when the EU had only 12 Member States have had to be updated (EU, 2016). In parallel, involvement of all animal stakeholders (other than authorities) in surveillance for so called regulated but also non-regulated diseases was shown to be of main importance for successful control of disease. In such a way,

several local and national initiatives emerged (Bisdorff et al., 2016; Calba et al., 2015, 2016; Cowie et al., 2015; More et al., 2015).

SURVEILANCE LEGAL BACKGROUND

International animal health standards

Following a rinderpest epidemic in Belgium, the OIE, formerly known as the '*Office International des Epizooties*' was founded in 1924 to coordinate disease control at an international level to avoid its spread across borders. After World War II, the Food and Agriculture Organisation (FAO, 1946) was established under the United Nations (UN, 1945), as well as the World Health Organisation (WHO, 1948). The General Agreement on Tariffs and Trade (GATT, first agreed in 1947 and finalised in Uruguay in 1994) included Sanitary and Phytosanitary (SPS) measures. This later gave rise to the World Trade Organisation (WTO, 1995). In 2003, the OIE was renamed to become the 'World Animal Health Organisation' but kept its acronym (Figure 1).

Figure 1 shows the main legal bodies that are involved in animal health surveillance and in drafting animal health regulation.

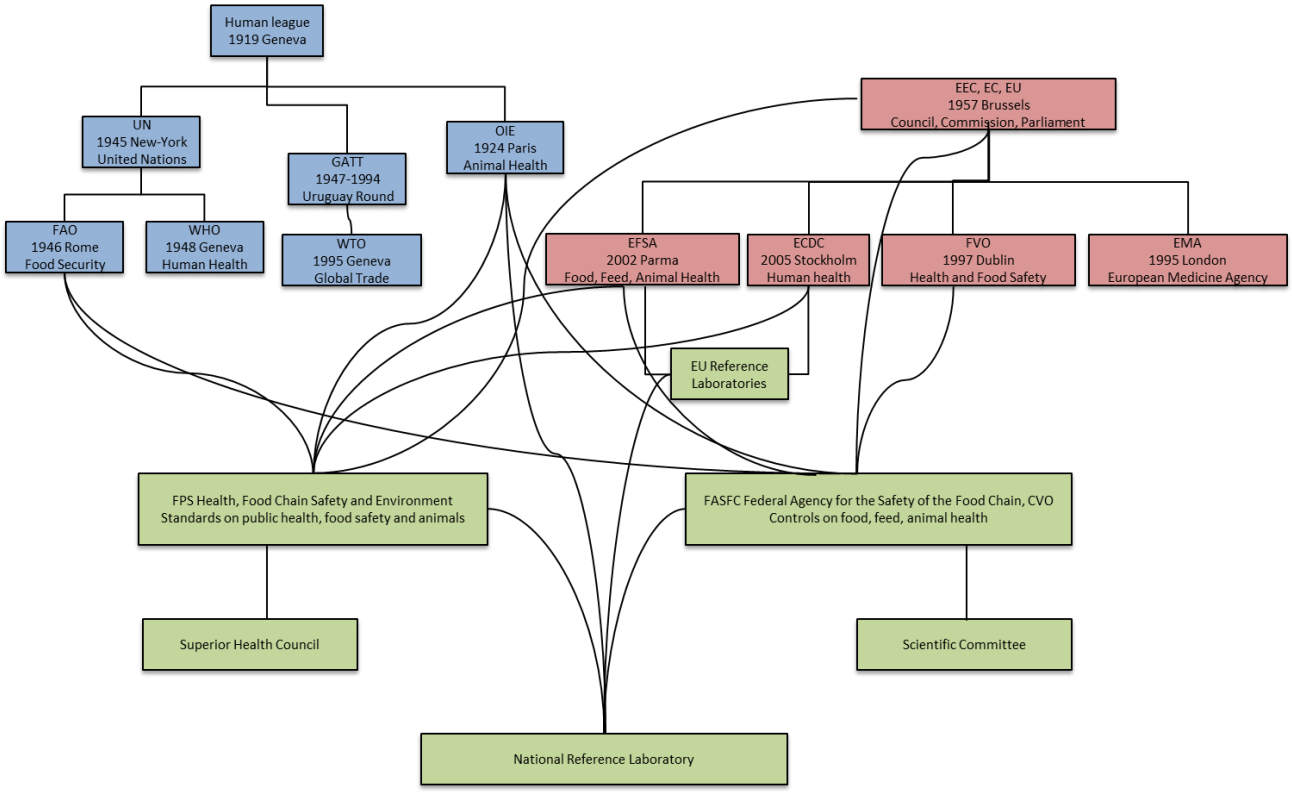


Figure 1: International, European, and national legal entities involved in animal health surveillance and in editing animal health legislation and standards: United Nations (UN), World Animal Health Organisation (OIE), General Agreement on Tariffs and Trade, Food and Agriculture Organisation (FAO), World Health Organisation (WHO), World Trade Organisation (WTO), European Economic Community (EEC), European Commission (EC), European Union (EU), European Food Safety Agency (EFSA), European Centre for Disease Prevention and Control (ECDC), Health Food Audits and Analysis (FVO), European Medicine Agency (EMA).

Following SPS measures, concepts such as equivalence, transparency, risk analysis and science based decision-making were introduced in order to mitigate risk of contracting infections through the movement of animals or products of animal origin, to ensure mutual trust and safe trading amongst partner countries. The OIE was later mandated by the WTO as the organisation responsible for drafting international standards for animal health. The tasks entrusted to the OIE are designed to:

- guarantee transparency on the status of diseases around the world;
- collect, assess and distribute information generated by veterinary science;
- mobilise expertise and stimulate international collaboration on the control of animal diseases;
- ensure safe trading in animals and animal products issuing health standards;

- promote legal frameworks and veterinary service resources at national levels;
- guarantee food safety and promote animal health following a scientific approach;
- provide science based guidelines and regulation for surveillance and control programmes, to ensure safe trading between countries.

All standards feature in the “Terrestrial Animal Health Code” and the “Manual of Diagnostics tests and vaccines for Terrestrial Animals” (as well as their equivalent for aquatic species) (OIE, 2017a,b).

In its Manual on Livestock Disease Surveillance and Information System together with its competencies on food and agriculture, the FAO also issues rules to consider while setting up and assessing surveillance programmes (FAO, 2017).

It must be stressed however that neither the OIE nor the FAO have legal authority, they only draft standards.

European animal health standards

In contrast to the OIE and FAO, the European Commission (EC) has legal authority and power to decide on what actions to be taken in case of deviation from baseline standards (which, in effect, derive from or are inspired by those laid down by the OIE (OIE, 2017a,b). There are three main institutions involved in EU legislation. European parliament (represents and is directly elected by EU citizens), Council of European Union (represents the Member State governments and the presidency is on a rotational basis), the EC (represents the interest of the Union as whole). The EC is supported by European agencies such as the European Food Safety Authority (EFSA), the European Medicine Agency (EMA), the Health Food Audits and Analysis (FVO) and the European Centre for Disease Prevention and Control (ECDC) for drafting laws and policies. EC proposes the new laws and policies and the parliament and council adopt them. It’s then EC and Member States that implement them and EC have the mandate to check that they are implemented and correctly applied.

The three main legally binding type of official documents from the EU, namely Regulations, Directives and Decisions differ in that (i) Regulations are similar to national legislation and apply directly as such within each Member State, (ii) Directives set the rules that must be incorporated into all national legislations, according to local situation, and (iii) a Decision addresses specific issues and might apply to member states or organisations that are mentioned specifically and is

directly applicable (EU, 2017). The current body of legally binding EU documents covering the food chain consists of nearly seventy pieces of legislation.

A few examples of different legislations ruling animal disease surveillance in EU, are mentioned here below, those are also the ones addressed and or used along this thesis:

Maybe the most pivotal official document ruling animal health is the directive 64/432/EEC dating from 1964. This regulation lays down the minimum requirements to meet for achieving bovine and swine disease-free status enabling free trade in animals and products between member states enjoying similar disease status (EC, 1964). More recently are 2010/367/EU and 456/2012/EU documents that apply to Avian Influenza and Bluetongue respectively (EC, 2010, 2012) in response to those new (re-)emerging epidemics. In addition, mandatory reporting by each Member State on surveillance results 92/117/EC revealed important issues regarding insufficient or incomplete data collection (EC, 1992). Consequently, directive 2003/99/EC emphasised the need to harmonise data collection to ensure better comparison of trends and sources of different zoonosis, allowing better prevention of risk and threat for public health and safe trading (EC, 2003a). Later, partly as a consequence of the BSE crisis, amongst others, the need for better traceability of animals and animal products triggered the emergence of various regulations and directives on the identification and registration of cattle and its products (i.e., 1760/2000/EC, 911/2004/EC, 1082/2003/EC and 644/2005/EC that rule on ear tag identification, the holding of registers and the issuing of passports, the minimum level of controls to be carried out in the framework of the system and the need to keep data) (EC 2000, 2003b, 2004, 2005). This in turn provides huge data sources that enable better risk assessment.

Faced with a constantly evolving legislation, EC decided to lay down a new animal health policy that would replace the existing series of interconnected policy actions by a single framework encompassing all aspects (not only the absence of disease, but also taking into account animal welfare, feed, food safety and public health).

Main issues of gaps in key areas and deliverables faced with the previous legislations were assessed in order to provide a modernised, simplified, more risked-based approach for the protection of health and to ensure the use of efficient control tools so that regulation guiding processing along the food chain is applied effectively.

This led to the new Animal Health Law: Regulation EU 2016/429 (31st March 2016), establishing a clear and transparent regulatory framework in order to prevent losses and reduce disease impacts on animal health, the economy and public health (EU, 2016). Through the years, a drift from input based standards towards output based standards is observed. Meaning by this more freedom but also more responsibility is given to Member States for setting up there guidelines and proofs of meeting minimum requirements.

National animal health standards and surveillance in Belgium

In Belgium, the Federal Public Service for Health, Safety of the Food Chain and Environment (FPS) together with the Federal Agency for Safety of the Food Chain (FASFC) and the sanitary fund have the mandate to draft relevant animal & plant health national regulation (FPS, 2017). This is done in collaboration with the involved stakeholders (farmers, traders, veterinarians' organisations, regional and national reference laboratories, slaughterhouses andfood business operators). These legislations cover a range of activities in addition to EU requirements. The mandate of the FASFC is also to control, ensuring that legislation is respected. These entities are inspired by the minimum requirements laid down in EU and International documents, as well as those emanating from the Scientific Committee hosted by FASFC. For surveillance of notifiable (but also non-notifiable) disease in Belgium, farmers and local veterinarians (practitioners) are at the frontline for the detection of symptoms and collect samples. If they see alarming symptoms they are obliged to report to official veterinarian authorities at the FASFC. The FPS co-operates with the FASFC and the representatives of farmers (AGROFRONT) for the regulated diseases. Regional animal health organisations (ARSIA and DGZ) as well as other private laboratories collect and analyse samples obtained by the practitioners according to a quality system (ISO 17025). They centralise results, communicate with the practitioners and farmers and they report at national level or send samples that need to be tested for confirmation to the National Reference Laboratory (CODA-CERVA). CODA-CERVA operates at the national level, organising, coordinating, centralising and analysing samples that require confirmation testing from all regional laboratories. The epidemiology, risk assessment and surveillance unit (ERASURV) at CODA-CERVA supports authorities (FASFC and FPS) for the organisation and assessment of national surveillance programmes and it provides epidemiological and

scientific background information. Exchanges between these partners are permanent, taking place during formal technical working groups by sector (cattle, poultry, pigs, small ruminants, milk) or by disease (Infectious Bovine Rhinotracheitis (IBR), Bovine Viral Diarrhoea (BVD), Bovine Tuberculosis (TBC), Salmonellosis, and Bovine Brucellosis …etc.) (Figure 2). In addition other private organisations (i.e. rendering plants, milk collection units) or community funded organisations (i.e. universities) are also involved in animal health surveillance and are of main importance, however for readability purpose, only those strongly involved in notifiable disease surveillance are mentioned here.

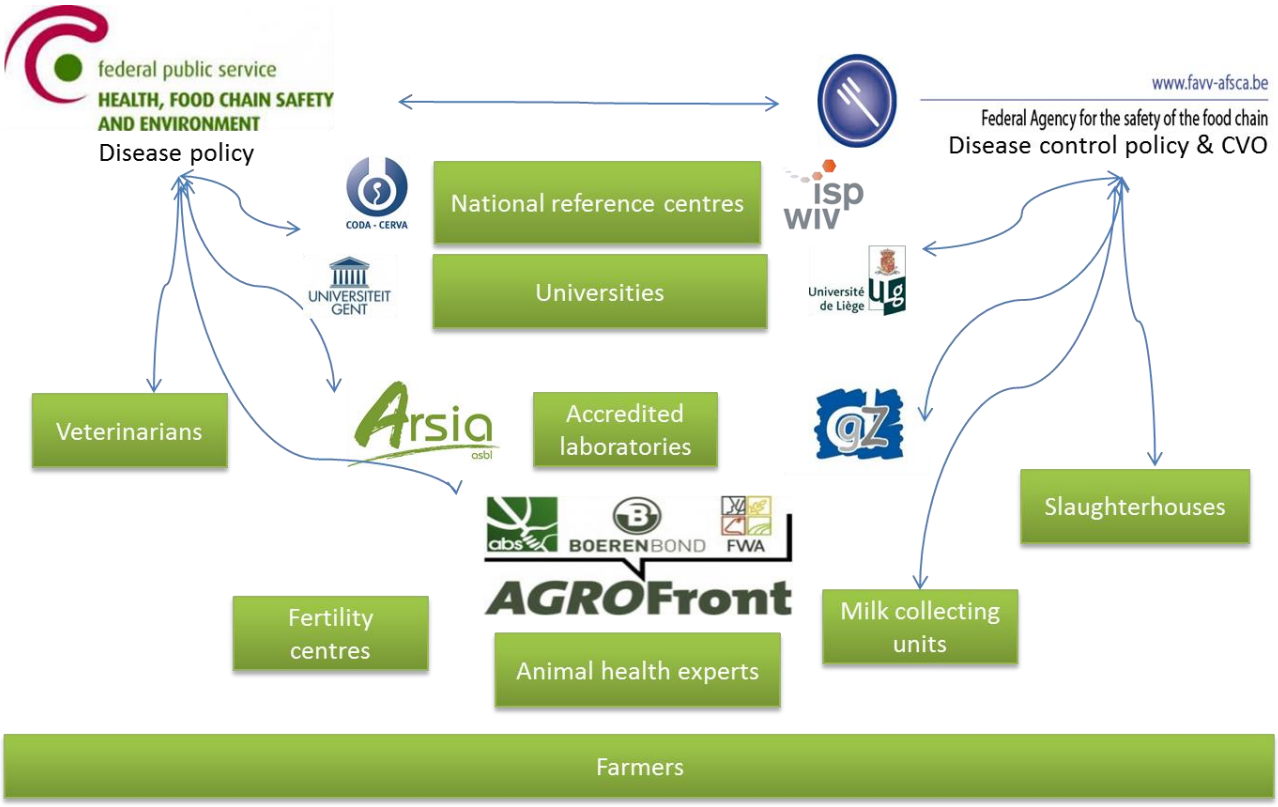


Figure 2: Main actors involved in the Belgian surveillance system for regulatory diseases (Federal Public Service for health, food safety and the environment (FPS), Federal Agency for the Safety of the Food Chain (FASFC), Veterinary and Agrochemical Research Centre (CODA-CERVA) Public Health Institute (ISP-WIV), Regional animal health organisations (DGZ, ARSIA), Universities, Veterinarians, Slaughterhouses, Millk collecting Units, Fertility Centres, Agriculture Farming Organisations, Farmers syndicate (AgroFront (ABS, Boerenbond, FWA).

Efficiency in tracing sources of infection or diseases and in implementing control measures will be driven by the quality of data (FAO, 2017; OIE, 2017a, b). To comply with the European legislation 1760/2000/EC, Belgium has created a database to gather information on animal identification and movements. The SANITEL database, a national computerised identification and registration system for all bovine, swine, ovine, caprine, cervids herds and poultry farms, contains for some species information regarding date and place of birth, previous and current place of holding, dates of on-farm and off-farm movements and date and reason of death. Although SANITEL has not been implemented specifically for disease surveillance, it is of high use as information regarding the disease status and certification of herds for official surveillance programmes such as Aujeszky, IBR, BVD, TBC, and Bovine Brucellosis. The day-to-day management of this national data base is entrusted to the FASFC though regional laboratories (ARSIA and DGZ) are at first line for entering the data. LIMS (Laboratory Integrated Management System) data are laboratory data obtained from all laboratories involved in national animal health monitoring and surveillance programmes in Belgium. Data is recorded at regional level by the different labs then centralised at the national level where responsibility lies for data validation, enabling later surveillance analysis and evaluation for the whole country by epidemiologists from the Veterinary and Public Health Institutes (WIV-ISP and CODA-CERVA). Farmers and veterinarians are at first line for collecting the data and ensuring the correct implementation of sanitary measures. In Belgium to encourage participation of farmers in surveillance, a sanitary fund system was created. This sanitary fund enables compensation in case of notifiable regulated disease outbreaks that requires specific measures, but also enables surveillance and control measures for non-notifiable disease. The interplay and involvement of all animal stakeholders (from farmers to authorities) challenged by regular working groups for each animal species and disease enables to take sound decisions and the implementation of sustainable surveillance systems and animal health in Belgium. However, tools to assess whether established standards are met or not or assess how confident one can be in surveillance are required.

THE ESSENTIAL STAGES AND FEATURES FOR EFFECTIVE SURVEILLANCE

Stages and features of surveillance

Containment and control of epidemic diseases depend on timely and good-quality information about disease events in order to understand the disease situation, support decision-making and prevent potential disease incursion. Various tools for collecting information on animal health at national and regional levels have emerged. However, to assess the drivers of animal disease and the patterns of transmission and spread, there remain challenges relating to the sensitivity of surveillance systems to capture information about new hazards or re-emerging threats. Epidemiologists and laboratories network play an important role in collecting, collating and analysing data on diseases and in providing epidemiological interpretation of the obtained results and converting these results into information to guide planning of disease control and risk management (Drewe et al., 2012; FAO, 2017). A brief guideline of key principles and strategic framework for effective surveillance is presented in Table 1.

Table 1: Strategic framework and key features for effective surveillance

Essential stages and features to consider in surveillance	
Stakeholders involved	• Decision makers (regional and/or federal), Authority (regional and/or federal), Laboratories (regional and/or federal), Universities, Field veterinarians, Farmers, Consumers, Food business Operators

Description and aims of surveillance	• Importance of the hazard: (frequency/severity/economic/zoonotic impact)
	• Objective:
	  ○ Baseline level: assess spread and impact of disease or intervention
	  ○ Case detection: (early-)detection of (re-)emerging disease
	  ○ Substantiate freedom of disease
	• Purpose (rules, education, social behaviour, ethical, animal welfare and tools)
	• Resources: time/funds/field labour
	• Attributes of surveillance to assess (sensitivity, specificity, representativeness, effectiveness, efficiency, timeliness...)

Surveillance components	• Passive, Active, Syndromic surveillance and ontologies
Surveillance Population	• Representative population, Target population, Sample frame, Study population

Surveillance design	• Sample size, Diagnostic tests characteristics, Sample frequency,
	• Sampling strategy:
	  • Non Probability sampling: judgement, convenience, purposive
	  • Probability sampling: simple random sampling, systematic random sampling, stratified random sampling, multistage sampling, cluster sampling, targeted hazard versus risk-based sampling

Stakeholders involved

Whereas public and animal health surveillance is a tool to assess the health status and the impact of measures implemented by authorities in charge of agriculture, health and finance—or other sponsors and donors — data from effective surveillance systems is also useful for these same bodies to target resources and assess programmes. Thus, legal entities (national governments, international authorities) certainly have a decisive role to play to ensure public health and to respond to threats. However, on the other hand, animal owners, the health industry and carers are probably better placed to mitigate risk. Stakeholders' needs and priorities might differ from those of authorities, but commitment and participation of all parties are essential for effective and successful surveillance (Calba et al., 2015; Dufour and Hendrikx, 2007; EC, 2014).

Description & aims of surveillance

Disease surveillance activities are designed to ascertain a given population's health status, for either early detection, or to assess prevalence and impact or to prove absence of animal diseases (Figure 3). Ultimately disease surveillance will inform and guide actions to its control (vaccination, eradication, or contingency planning). The design will be tailored to the goals and disease status. However, economic, political, social and welfare aspects can also determine surveillance activities. The various stakeholders in surveillance can sometimes contribute to the apparent inconsistency of surveillance goals since they can vary according to stakeholders' priorities and needs (Cowie et al, 2015; Stark & Häsler, 2015).

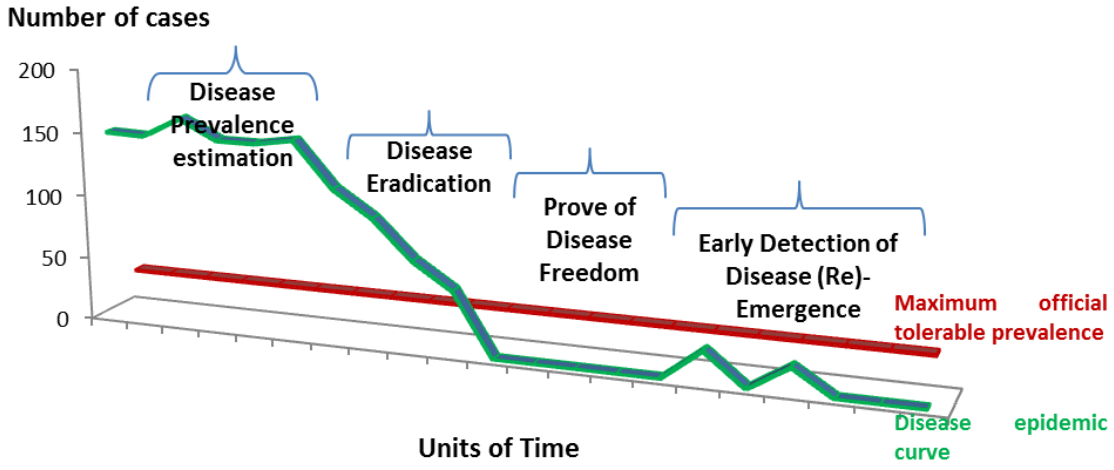


Figure 3: Surveillance based on disease status and goals. While disease is endemic,

the aims are to assess prevalence, deviation from baseline trends, impact assessment (e.g. vaccination coverage) and eventually eradication. Later, once disease has been successfully eradicated early detection of emergence or re-emergence of disease together with proving absence below the maximum tolerable prevalence will be the main aims guiding surveillance activities.

Although surveillance systems may vary widely in methodology, scope and objectives, such variability can be harnessed by encouraging the definition of clear objectives and purpose, and by reporting according to a standardised terminology (Bisdorff et al., 2016; Drewe et al., 2012; Calba et al., 2015; Hoinville et al., 2013).

To ensure surveillance meets desired objectives, several attributes are mentioned throughout the literature and official standards to implement and assess surveillance (i.e. objectivity, acceptability, accuracy, transparency, data quality, efficacy, efficiency, feasibility flexibility, positive predictive value, negative predictive value, relevance, representativeness, security, sensitivity, specificity, simplicity, stability, timeliness, usefulness) (Cameron, 2012; Drewe et al., 2012; Hoinville et al., 2013). It is recommended to not only rely on one or two attributes to set up or assess surveillance (Drewe et al., 2012), but efforts to improve certain attributes — for instance the ability of a system to detect a health event (sensitivity) might be detrimental for other attributes, such as simplicity or timeliness. Therefore success of a surveillance system will depend on the right balance of characteristics on the one hand and, on the other, the ability to adapt these characteristics according to the system's requirements and resources (Drewe et al., 2012).

Surveillance components

A wealth of necessary information to assess surveillance can be gathered from data collected by farmers, slaughterhouses, laboratories, dairy collection units. Each separate flow or source of data generated defines the components of surveillance. The wide variety of data collection sources, storage format and flows of information obviously make management more complex but they can be summarised as three main categories of surveillance component: passive surveillance components (such as clinical or abortion notifications), active surveillance components (such as sentinels or repeated cross-sectional serological surveys) and syndromic surveillance components (defined by the collection and analysis of non-disease-specific

production data (i.e. milk production data; reproductive performance data; mortality data) to detect deviations of the baseline trends).

More specifically, from an animal disease angle or at a national level, surveillance can be seen as a system comprising several surveillance components (defined as surveillance portfolio) which are characterised by their specific population coverage and data processing system, each of them contributing to inform on surveillance objectives and targets; i.e. bulk milk testing will only cover dairy herds and data generated will be recorded separately; slaughterhouse surveillance will cover a large population and, consequently, the flow of recorded data will be large too and belong to a separate stream (Hoinville et al., 2013). Surveillance components though mainly described as active or passive surveillance are occasionally described as respectively proactive or reactive (Dufour and Hendrikx, 2007), however along this thesis we will stick to the two former. Passive surveillance relies on notification and goodwill from farmers/veterinarians only, who report/notify suspicious animals. This can be considered as a frontline tool, where farmers and animal care workers, in close contact with animals daily, are relied upon to provide a constant and immediate picture of the disease status in a given population. Fear of ethical repercussions or economic repression, visibility of clinical signs, and the awareness of farmers and animal care workers are the main constraints that hamper the efficacy of such surveillance components (ARSIA, 2013; Dufour and Hendrikx, 2007; Elbers et al., 2010; Humblet et al., 2011). While public health surveillance mainly relies on passive surveillance, animal health surveillance, to ensure food safety or fulfil trade requirements, triggers the need of sufficient data and its representative collection via active surveillance. Active surveillance implies testing individuals from a sample of the target population, regardless of whether or not they show clinical signs. It provides a representative and/or snapshot image of the disease status in the population or it detects sero-conversion through the routine serological testing of sero-negative individuals (Racloz et al., 2006). The fact of planning limits detection to predetermined indicators, thereby hampering early detection of an unexpected (re)-emerging disease. Syndromic surveillance makes use of non-specific data on animal health, production or mortality. Investigation of routinely collected data (i.e. milk production data, reproductive performance data), or population levels (laboratory

data) enables the assessment of deviations from baseline trends. Over recent years, this form of surveillance has gained popularity (Hoinville et al., 2013).

Surveillance population & surveillance design

Because census sampling is not feasible for practical and economic reasons, inference towards the target population is made on its disease status from a sample of the reference population. Sampling can hamper accuracy and precision of the population estimates such as totals and means. This is particularly the case when the study population (selection from the sample frame (fraction of population from which we can draw a sample)) characteristics differ from the target population (one for which surveillance is meant for) (Dohoo et al., 2012; Molenberghs, 2009).

Lack of accuracy is called bias. Bias is the deviation of the sample estimate from the true population value and is the result of a systematic error. Bias can be caused by different aspects of the study's design, such as: 1) the sampling (selection bias); 2) diagnostic test characteristics sensitivity and specificity (measurement bias) or 3) the method of analysis (analysis bias). Precision is the estimate of the variation and uncertainty around an estimate, and it is the result of random error. The (lack of) precision is affected by the number of samples and variance in population. Because of bias and precision uncertainty, the inference for the obtained statistics (population estimates) can result in errors. Each inference is accompanied by probability errors. Type I error (alpha), which is concluding an effect exists when in fact there is no effect. In surveillance, type I (alpha) error is related to the herd specificity (Alpha=1-herd specificity), which indicates the false positive rate. Type II error (beta), which is concluding there is no effect or differences when in fact there is. In surveillance this is related to the herd sensitivity (Beta=1- herd sensitivity), which indicates the false negative rate.

Because tests are imperfect, a combination of different tests or components will be considered to increase the confidence about sensitivity or specificity of detection. Sensitivity, specificity, positive predictive value, negative predictive value, repeatability and reproducibility constitute the main parameters used to evaluate diagnostic test (Banoo et al., 2010).

Frequency of sampling can be adapted depending on disease characteristics (epidemic or endemic disease, slow or fast spreading, incubation or latency periods). The reproductive ratio (R0), defined as the number of secondary cases caused by

one typically infectious case (Velthuis et al, 2007; Graat et al., 2001), enables prediction of likelihood of major or minor outbreaks. An infectious disease with an R0 below 1 will eventually die out without intervention, whereas an infection with an R0 above 1 will likely spread and require intervention strategies to stop the spread of the disease. The likelihood of a major or minor epidemic, can be estimated as 1-1/R0, e.g. an infection with an R0 of 5 will have a probability of 80% to cause a large outbreak. By contrast, an endemic situation is characterised by diseases that are present within the country at low or high prevalence but that do not have a “sudden” onset in space or time (Buehler, 1998; Comin et al., 2012).

As the sampling design concerns, many possibilities exist. Probability sampling or non-probability sampling are the main sampling techniques. Non probability sampling, such as convenience sampling (i.e, holdings close to each other or all dairy herds), purposive sampling (holdings with a known exposure to the risk factor or having a specific disease status (i.e. only Salmonella positive holdings) or judgment sampling (based on researchers decision or human decision only) are easy to carry out. Yet, their disadvantage is that they are likely to introduce selection bias. Probability sampling has the advantage of allowing inference to the population. Simple random sampling implies that individuals have non-zero and equal probability of being sampled. This method allows inference to the whole population taking into account sampling weight (one only selects a fraction of all existing holdings). In systematic sampling, a fixed interval is used to decide what individuals to select from a list of the population to be sampled. Stratified and proportional to size sampling, allows greater precision of the overall estimates. Usually, stratified sampling is carried out as a proportion of the size of a stratum (i.e., number of herds per provinces or regions). Hence, small size region are sure to be selected, which would not have been the case if done by a simple random sample. However, larger strata will then have less weight in the overall estimates, as would have been the case in simple random sampling. Also, depending on sample size allocation per strata, the stratum specific estimates might have lower precision. When clustering effects are ignored (i.e. two pairs of overshoes of the same farm will be correlated), there will be lack of accuracy and precision. Multistage selections of holdings and of animals takes into account the clustering effect. The variation within and between herds must be known to estimate the correct sample size. The sample size required will be larger

for a same desired precision. Clustered sampling is based on the principle that selecting holdings—and within those holdings sampling all animals—is easier to carry out, as often the list of holdings exists, which is not the case for animal IDs. It is also cheaper to sample all individuals from a small selection of herds than it is from a selection of individuals from a large number of herds. However, due to the correlated structure of the data for a given sample size (animals from the same farm are more likely to behave the same way or have the same disease status); the variation within herds will be small whereas variation between herds is larger. Depending on the correlation within a herd— in particular when within-herd correlation is high — it is better to sample many different herds than many animals within a limited number of herds in order to get a precise and unbiased estimate of disease status prevalence (Faes et al., 2011).

Targeted sampling can be considered as a special case of cluster sampling and it can also be referred to as risk based surveillance, where the aim is to detect specific hazards that are more likely to occur in specific strata of the population. Despite increasing the efficacy of sampling and reducing field work, precision and accuracy of estimates will be altered and need for correcting on the basis of within and between classes correlations will be generated. Risk based surveillance offers a very useful alternative to improve efficiency under the following conditions: epidemiological expertise, knowledge regarding target population, good data quality, and transparency. Only under these conditions, risk based surveillance can outperform a random sampling to detect disease or follow up the implemented measures. Combining different sources of evidence is also an alternative to improve efficacy of surveillance as well as accounting for historical data (Bisdorff et al., 2016; Cameron 2012; EU, 2016). This automatically triggers the need of tools to correct the precision and biased nature of the sampling in order to allow inference from data generated by the surveillance system (Dohoo et al., 2012; Molenberghs, 2009). More attention to this topic will be addressed further in the thesis.

SURVEILLANCE EVALUATION

Models and simulations in surveillance

Where time and budgetary restrictions prevent the conduct of field experiments, simulation models may offer good alternatives. To model disease processes in

animal populations or to quantify the impact of measures from standard risk factor analysis by means of regression techniques to simulate spreading (simulating outbreaks or estimating reproductive ratio), simulation models can provide better prediction than what one can simply observe in an experimental unit. Models can be built from field data in order to realistically simulate real life scenarios.

Simulation models can be deterministic or stochastic. Deterministic models are relatively easy to implement because they simulate results of a given scenario for a set of predetermined values and returns one single outcome value. However, conclusions from such models do not take into account the possible variation and uncertainty of events. In contrast, stochastic models simulate events as a function of a range of different values. Biological variability, complexity of disease transmission, uncertainty around current knowledge and a country's disease situation can all be factored in by means of probability density functions and reflected in the output characterised by an outcome distribution. The appropriate probability density function or distributions will be defined by the probability that a random variable will fall within a particular range of values. The parameters will describe the shape of the probability density function. For instance a normal distribution will be characterised by the mean and the variance. The binomial probability distribution is often used to characterise diagnostic tests, and is characterised by the parameters n (number of trials) and p (probability of success). The pert distribution often used in stochastic simulation models, and used to describe variability and uncertainty originating from expert opinion. It is characterised by the parameters minimum, maximum and most likely value. These same parameters are used to define the triangular distribution, however, in contrast to the triangular distribution; the pert distribution will give more emphasis to the tails of the distribution. The combination of all probability density functions will generate the output probability distribution.

Common processes for stochastic simulation are the "Monte Carlo" algorithms but others exist as well such as the "Latin hypercube" sampling amongst others. The differences lie in the iteration processes. Whereas "Monte Carlo" will randomly pick up random values out of the input distribution, the "Latin hypercube" sampling process will pick up more values in the mean of the distribution and it places less emphasis on the tail of the distribution. The output of a stochastic simulation model can be seen as a probability density function, reflecting the uncertainty and the

variability of possible outcomes of a given scenario. Stochastic modelling has gained increased recognition in veterinary epidemiology. However, one must take care when interpreting results from such models. Indeed, one of the pitfalls of these models is that outputs are strongly dependent on inputs; the “NINO” concept (Nonsense In, Nonsense Out) is sometimes used to describe this pitfall. Therefore, validation of the model outcome as well as the used probability density functions, is an important step. One of the fields where simulation techniques have gained considerable interest is in the assessment of probability of freedom from a specific disease. In other words, given that no positive reactors are observed, what is one's level of confidence in true freedom, taking into account the surveillance system's design (Boklund, et al., 2013; Frossling et al., 2013; Knight-Johnes et al, 2010). Rather than evaluating input based standards, one sets the output based standards (i.e. confidence level in freedom of disease in a surveillance setting of a certain disease with low prevalence). Simulations are then used to ‘mimic’ the sampling process within certain target/risk groups as well as ‘mimicking’ the combinations of sampling processes in the different risk groups. It allows assessing which sampling process might be the most successful taking into account the different surveillance attributes described above in Table 1. In addition, cumulative confidence gained through time with the disease surveillance process can be modelled using Bayesian theorems.

Assessment of surveillance effectiveness and efficiency

Obviously, it is of the highest importance that surveillance should be capable of achieving its objectives; this is what is commonly defined as effectiveness. Whereas it is clear that in a context of emerging and remerging diseases there is a need for effective and cost-efficient surveillance systems (as much for public health as for animal welfare and trade), criteria and tools that can provide mutual trust and sufficient confidence about the country’s surveillance are still lacking (Bisdorff et al., 2016; Calba et al., 2015; Drewe et al.; 2012; Hoinville et al., 2013; Stärk and Häsler, 2015). Stark et al. (2002) and Salman et al., (2003) introduced the need of quality assurance and assessing surveillance. Later, Martin et al. (2007) and Cameron (2012) summarised the evolution from input based standards to output based standards, the latter allowing better comparison of quality equivalence. Furthermore, striking a balance between benefits from a surveillance system and the cost of running it, in other words surveillance efficiency, is important too. It is important to

consider the attributes for evaluation in relation to the objectives of surveillance (Drewe et al., 2012; Calba et al., 2015). Studies used to assess surveillance systems can either be quantitative or qualitative. In Table 2, all methods used in the context of animal health are summarised.

Amongst qualitative methods, the most common one is **Expert Knowledge Elicitation (EKE)**. This method uses a scoring system from experts who assess surveillance components. These methods were widely used in the past to forecast disasters and were later applied to other fields of risk assessment. The drawback of the method comes from two main assumptions: (i) subjectivity around scoring and (ii) group judgement is given more importance than individual judgement. Different techniques can counter these assumptions such as different weighing of the scores per expert — based on expertise — and reducing variability by repeating the process until consensus is reached. Logic models are an expanded subset of these tools, including all aspects of surveillance under consideration. In addition to scores, comments can be given too. Not only experts can be considered but all stakeholders in animal health surveillance. Consensus on scores is less important, the emphasis is placed on detecting similarities and differences in surveillance priorities, gaps and needs and present output visually by different means, which allows the capture of socio economic and cultural aspects (Calba et al., 2014, 2016). Also, in contrast to "top down evaluations" (i.e., where surveillance evaluation is carried out from authorities' angle and objectives essentially), it will ensure that points of view from the field as well as their expectations and perceptions are included in a "bottom up evaluation" and it can also bring together stakeholders, authorities, decision makers, program funders and beneficiaries (Calba et al., 2014, 2016; Queenan et al., 2016). However, these **participative and social epidemiology** methods require a representative number of stakeholders to sample, it will also require time to schedule interviews separately, in order to avoid bias in answers (interviewing authorities together with a community of farmers could change perceptions and answers). Finally, quantitative output results are hard to obtain as a scoring system remains subjective. This subjectivity around the output means that evaluation and validation of methods will be difficult to measure and, in the field, applying recommendations from lessons learned will be hard too.

Amongst the quantitative methods, **case reports** generated from data collection and summarising information obtained from different data collection streams is the primary step in surveillance. This can be done by simply a proportionate comparison of incidence or prevalence rates obtained either from surveillance components or from different regions (Bronner et al., 2014).

Although simple, common **regression techniques** (logistic, linear, longitudinal…) can be adapted and encompass a larger range of evaluation techniques than a simple risk factor analysis (Méroc et al., 2012). Timeliness, compliance and coverage can be assessed taking into account random effects linked to veterinary practitioners and other factors; and provide reliable information on solutions that could be implemented to increase compliance and timeliness and thus offer promising options for better surveillance (Del Rocio et al., 2010). Also, in terms of timeliness and geographical coverage, correlation between data collected from different surveillance systems and national or global data bases, provides useful information on the quality of a surveillance system (Farnsworth et al., 2010). For zoonotic diseases, extending comparison of animal disease surveillance components with human outbreak data provides interesting insight on transmission characteristics and predicting disease outbreaks, but also can help to identify important gaps in surveillance (Rabinowitz et al., 2012).

After identifying gaps, measuring the effect of efforts made is an important element to assess the quality of surveillance (for instance before and after stimulation campaigns and understanding reasons and motifs why cases are reported or not) (Elbers et al., 2010; Humblet et al., 2011).

However, the evaluation of the quality and timeliness of data is more complex. **Simulating epidemics** and their spread in space and time based on spread parameters, risk factors and intervention measures enable the evaluation of how well the intervention or surveillance protocols are capable of reducing the size of the final epidemic. Reproductive ratios, movement patterns, network analysis and spatial clustering can be used to simulate the spread of a specific disease and assess the impact of control measures (such as bans on movements), and efficacy of diagnostic testing protocols; or, alternatively, if different strains were to emerge with different reproductive ratios, they can assess their impact on surveillance (de Koeijer et al., 2011; Ensoy et al, 2013, 2014; Hasala and Boklund, 2014).

Syndromic surveillance, based on routinely collected production data (milk or fertility), mortality data and laboratory data provides an insight into the impact of certain diseases and their spread in space and time, which also provides valuable information to quantify efficacy of surveillance compared to conventional surveillance, (Dórea et al., 2013; Marceau et al., 2014; Perrin et al., 2012; Welby et al., 2016).

Capture recapture methods, initially developed and widely used for environmental surveys, are methods designed to estimate true population sizes mainly but they can be used in epidemiology to measure the thoroughness of disease records (i.e., quantify expected abortions and compare with reported abortions (Bronner et al., 2013). However, these methods are subject to constraints due to the assumptions on which these methods rely. The assumptions are that every animal in the population has the same probability of capture each time it is sampled; the probability of recapture does not change with time (Vergne et al., 2014).

Cost benefit analysis provides valuable information to assess surveillance efficiency. In addition to all techniques developed in this section, the principle here is to compare the benefits gained from surveillance (i.e., sensitivity) with efforts made; field requirements and budgetary burden. However, major limitations of this method are the value of information and inputs. Direct costs may easily be estimated but indirect costs or benefits may be difficult to estimate; i.e., what is the cost if surveillance is less sensitive, leading thereby to disease spread and production losses. The subjectivity around cost estimates, the different weights given to expected costs depending on stakeholders' views together with the lack of standardisation can explain the limited number of available studies for animal disease surveillance (Häsler et al., 2012; Drewe et al., 2014; Stärk and Häsler, 2015).

Output based methods, which will be developed in detail in the next section, is an example of a method used for evaluation of surveillance. Ultimately it is conceived to measure whether surveillance meets the design targets. Different surveillance components can be compared with each other based on their relative efficacy (sensitivity) and the surveillance system as a whole; taking into account all information provided by the various components can be estimated too (Boklund, et al., 2013; Frossling et al., 2013; Knight-Johnes et al, 2010).

Table 2: Various quantitative and qualitative methods to assess the effectiveness and efficiency of surveillance systems (adapted from Drewe et al., 2012)

Methods	***Description of the method***	***Reference***
Participatory epidemiology Social epidemiology Logic model	Assess qualitatively and score experts and animal health stakeholders' priorities, gaps and needs for surveillance. The output is then visualised with different graphs, which enables identification of gaps and priorities in surveillance	Calba et al., 2014, 2016; Hendricks et al., 2011 ; Queenan et al., 2016; Roelandt et al., 2015
Case reports	Assess the incidence or prevalence from different components or measurements of increased notifications before and after disease awareness campaigns	Alfonso et al., 2014 ; Bronner et al., 2014; Humblet et al., 2011; Elbers et al., 2010
Regression techniques	Odds of detection, relative risk of detection, correlation between different surveillance components data, compliance, timeliness of surveillance methods	Del Rocio et al., 2010; Farnsworth et al., 2010; Méroc et al., 2012; Rabinowitz et al., 2012; Welby et al., 2011
Spatial and spread models	Simulating and assessing impact of different surveillance and intervention strategies on spread or transmission of disease	Claes et al., 2013; de Koeijer et al., 2011; Ensoy et al., 2013, 2014; Hasala and Boklund, 2014
Syndromic surveillance	Identification of deviation from baseline trends	Dórea et al., 2013; Marceau et al., 2014 ; Perrin et al., 2012
Capture, recaptures models	Resampling method	Bronner et al., 2013 ; Vergne et al., 2014
Cost Benefit	Estimates cost in comparison with gained benefits; (i.e.,. sensitivity of detection)	Häsler et al., 2012; Drewe et al., 2014; Rutten et al., 2012; Stärk and Häsler, 2015
Output based methods	Modelling disease surveillance process quantitatively within the different populations of interest taking into account their differential risk of infection or detection enables assessment of the effectiveness of surveillance in relation to prescribed minimal standards	Boklund, et al., 2013; Cameron, 2012; Frossling et al., 2013; Knight-Johnes et al, 2010; Welby et al., 2010, 2012, 2013, 2016

Output based simulation models for substantiating freedom of disease

Traditional sample size calculation systems, developed by Canon and Roe (1982), have shown their limitations as they assume perfect test sensitivity and specificity. More generally tests used do not have these perfect test characteristics and survey results deliver estimates with a much lower confidence than expected. Cameron and Baldock (1998a, b), as well as Jordan and McEwen (1998), adapted these formulas to include herd level uncertainty and imperfect test characteristics, but they assumed representative sampling (Cameron, 2012). Doherr and Audigé (2001) introduced the need to combine data from different sources when claiming freedom of disease. They also underlined the need to demonstrate freedom whilst enabling detection too. Later, scenario tree, output based methods, introduced by Martin et al. (2007a, b) provided tools to account for non-representative sampling and differential risk of infection. Martin et al. (2007a, b) also enabled to account and combine data obtained from different surveillance components and data processes in order to indicate surveillance system sensitivity. In addition Cameron (2012) provided tools to account for historical information regarding surveillance over time.

The principles of such risk based scenario tree methods lie on the fact that one has more chances of detecting the disease if one has prior knowledge of where it clusters. By targeting surveillance in populations where disease tends to cluster, one can thereby increase efficacy (sensitivity). Each surveillance component activity is summarised into a bifurcation chart called scenario tree. For instance, components in Belgium that can be considered as pillars of Brucellosis surveillance in cattle are: abortion notification, purchase tests, import tests, bulk milk testing of dairy herds, tracing herds after breakouts over five consecutive years as well as breakouts from herd contacts showing equivocal results and, finally, yearly random sample. Figure 4 shows three scenario trees for 3 different surveillance components (abortion testing, testing cattle during ‘intra-country purchase’ and bulk milk testing of dairy cattle). Populations under review for these three components are different. The abortion component will concern female cattle essentially, purchase component will cover all cattle moved from one farm to another following purchase and bullk milk testing will concern dairy cattle in large enough dairy herds delivering milk via tank milk to milk collecting units. In spite of fairly distinctive population coverage, overlaps exist. Methods developed by Martin et al.. (2007a, b) allow adapting for the overlap.

However, it is quite common not to take this overlap into account, as in practice information generated by one component will not impact sampling in other surveillance components.

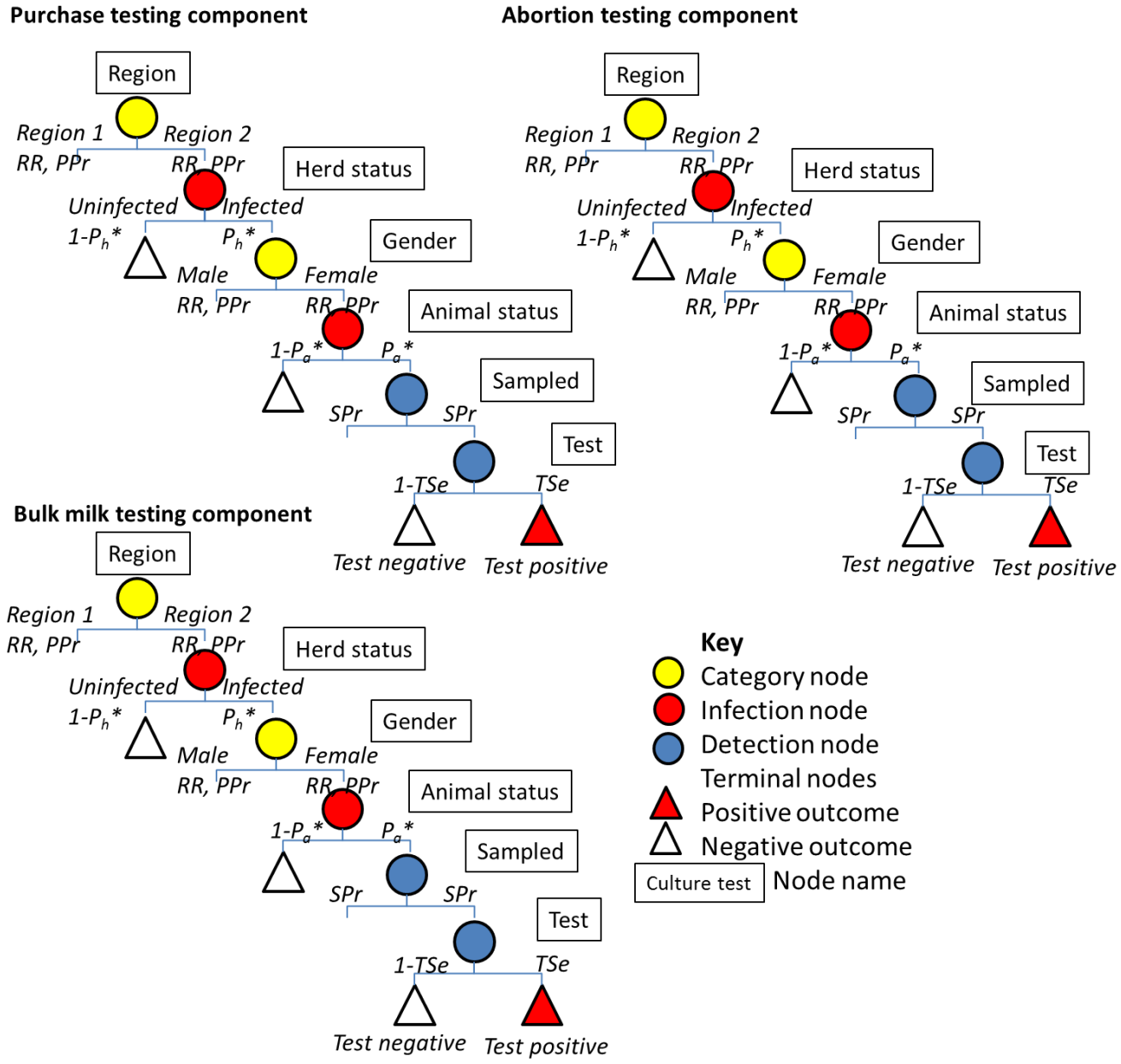


Figure 4: Scenario tree developed to evaluate the performance of different surveillance components (abortion testing, testing cattle during 'intra-country purchase' and bulk milk testing of dairy cattle herds) for detection of bovine brucellosis in Belgium. PPr, SPr are the relative population and sampled proportions in each risk group and RR are the relative risk of infection in each risk group. P is design prevalence at herd (h) and animal leve (a) while TSe is the diagnostic process sensitivity.*

The scenario tree is characterised by its nodes describing each event from infection to detection. Each node splits the population sampled (n) and the representative population (N) according to the branches representing the differential probabilities of infection or detection. Three types of nodes are considered: risk category nodes, infection nodes and detection nodes. They are placed in a chronological order (representing the sequence of events from infection to detection), according to decreasing size. The design prevalence (P*), at herd (h) or animal level (a), used for the model design differs from the actual prevalence and enables drawing conclusions on surveillance effectiveness of the component against an agreed standard. It is a set benchmarking value used to estimate the effective probability of infection (EPI) based on the relative risk (RR) of infection and population proportion (PPr) for each subgroup.

$$AR_i = \frac{RR_i}{\sum_{i=1}^{n} RR_i * PPr_i}$$

For each risk group i, the effective probability of infection (EPI) for holdings and animals is calculated as follows:

$$EPI = AR_i * P^*$$

The combination of each branch and node constitute the limbs of the tree, defining the subpopulation groups and their respective relative effective probability of infection and detection. Too many details about different components and their particular data collection processes tend to distance the output from the main issues to be addressed. Therefore, it is crucial that during this phase only major components are described. Only if sampled populations under review and the diagnostic processes are different should a new component be considered. Whereas each component can be evaluated and described separately, an overall output can be obtained for the whole surveillance system too.

The effective probability of detection (EPD) is the result of the whole diagnostic process in the given component (i) (probability of clinical signs or diagnostic test sensitivity (TSe)). The surveillance sensitivity of the component under consideration (CSe) is then obtained using the following formula:

$$\mathrm{CSe}_i = 1 - (\prod (1 - \mathrm{EPD}_i (\frac{\mathrm{n}_i}{\mathrm{N}_i}))^{\mathrm{N} * \mathrm{EPI}_i})$$

Where EPI is the probability of infection; EPD the probability of detection of an infected/affected unit after going through each of the stages to consider within the

specific limb of the given component; n is the number of units processed through that limb of N population size. For each surveillance component, including its own specificities and data collection process, a new and separate tree is considered.

To estimate the overall sensitivity of the surveillance system components (SSe), the following equation is used:

$$\mathrm{SSe} = 1 - (\prod(1 - \mathrm{CSe_i}))$$

In order to design a tree, preliminary knowledge on risk factors affecting the probability of detection or infection must be known. Assumptions can be used in the absence of existing knowledge. Simulations can be carried out when no knowledge exists. However, one must then make sure that validation is carried out for those assumptions. This can be done by internal and external validation (Gustafson et al., 2013). Both differ in the way they are carried out. Most of the known software use internal validation. In other words, the sensitivity analysis will mainly be affected by the uncertainty and variability around inputs within the model and those factors will be the most decisive. External validation of the model's robustness for instance can be done by testing the assumptions by setting some parameters in the model and changing others. External validation can also be done by confronting outputs of the simulation model and inputs with different inputs obtained from field data, or using data originating from a different country or a different component. Alternatively, validation can be conducted using an empirical model to test some outputs.

However, the question “how can one be sure that disease is truly absent when all samples test negative?” will still remain; even if the surveillance system and different components reveal themselves as being effective. This is where methods introduced by Martin et al. (2007a, b) may help by using a Bayesian approach to the problem. The prior knowledge of confidence in freedom, obtained from the surveillance, is taken into account for calculating the posterior probability in freedom. Thus gathering an accumulation of evidence over time ultimately gives you more and more evidence of disease status knowledge and the probability of freedom. These approaches also allow simulating various scenarios and measuring how well the infection could be detected given current surveillance systems.

$$\mathrm{Pfree_{t_i}} = \frac{\mathrm{1\text{-}Pinf_{ti\text{-}1}}}{\mathrm{1\text{-}Pinf_{t_i\text{-}1}{}^{*}CSe_{t_i\text{-}1}}}$$

$$Pinf_{t_i} = \left(1 - Pfree_{t_i - 1}\right) + Pintro_{t_i} - \left(Pintro_{t_i} * \left(1 - Pfree_{t_i - 1}\right)\right)$$

Where the posterior probability of freedom (Pfree) is obtained given prior probability of infection (Pinf) of the previous year and the surveillance component sensitivity (CSe) of the component under consideration (i) of the previous year. The Pinf can be determined arbitrarly, ie the expected design prevalence, the very first year. Then, for the following years, it will be function of $1 - \mathrm{Pfree}_{t_i - 1}$ and the probability of introduction (Pintro), which can also be set arbitrarly. One assumption deems that if a country is free of disease, any positive reaction will be further investigated until truly positive or negative. Therefore, specificity is considered to be 100%. In practice of course, this is not the case. However, this will not hamper the output as the question addressed by these methods is: "given one does not observe any positive (meaning truly positive after all confirmation tests have been carried out (serial testing until specificity nears 1), how confident can one be of enjoying a disease freedom status?"
Distributions are fitted to most of the input parameters allowing **stochastic simulation modelling.** The advantage of this method is that all quantitative parameters and stages in the disease detection process can be summarised, from expected prevalence, differential risk of infection or detection, test characteristics (sensitivity and specificity) to representative population proportion sampled. In addition, the advantage of these methods is to allow combining evidence of disease status from different sources or components when evaluating the combined national surveillance system (i.e., slaughterhouse data, bulk milk data, purchase testing data, export testing data). The relative efficacy of each separate component can also be estimated.
Also, the relative costs of surveillance within the different evaluated surveillance components can be computed and thus the cost efficiency ratio of surveillance can be obtained using the same methodology, as shown in Welby et al. (2016).

REFERENCES

Afonso A, Abrahantes JC, Conraths F, Veldhuis A, Elbers A, Roberts H, Van der Stede Y, Méroc E, Gache K, Richardson J, 2014. The Schmallenberg virus epidemic in Europe-2011-2013. Prev Vet Med. 116(4):391-403.

Association Régionale de Santé et Identification Animale(ARSIA), 2013. URL (Last update date : 01/2013) http://www.arsia.be/wp-content/uploads/2013/01/flyer-Avo-libramont2012-pixart.pdf (Consulted date: 05/12/2014).

Banoo S, Bell D, Bossuyt P, Herring A, Mabey D, Poole F, Smith PG, Sriram N, Wongsrichanalai C, Linke R, O'Brien R, Perkins M, Cunningham J, Matsoso P, Nathanson CM, Olliaro P, Peeling RW, Ramsay A, 2010. Evaluation of diagnostic tests for infectious diseases: general principles. Nat Rev Microbiol. 12 Suppl:S17-29.

Bisdorff B, Schauer B, Rodriguez-Prieto V, Comin A, Brouwer A , Dorea F, Drewe J, Hoinville L, Lindberg A, Martinez-Aviles M, Martinez-Lopez B, Peyre M, Pinto Ferreira J, Rushton J, Van Schaik G, Stark KD, Staubach C, Vicente-Rubiano M, Witteveen G, Pfeiffer D, Hasler B, 2016. Active animal health surveillance in European Union Member States: gaps and opportunities. Epidemiol Inf , pp. 1–16. doi: 10.1017/S0950268816002697.

Boklund A, Dahl J, Alban L, 2013. Assessment of confidence in freedom from Aujeszky's disease and classical swine fever in Danish pigs based on serological sampling--effect of reducing the number of samples. Prev Vet Med. 110(2):214-22.

Bronner A, Hénaux V, Vergne T, Vinard JL, Morignat E, Hendrikx P, Calavas D, Gay E, 2013. Assessing the mandatory bovine abortion notification system in France using unilist capture-recapture approach. PLoS One. 8(5):e63246.

Bronner A, Hénaux V, Fortané N, Hendrikx P, Calavas D, 2014. Why do farmers and veterinarians not report all bovine abortions, as requested by the clinical brucellosis surveillance system in France? BMC Vet Res. 10:93.

Buehler JW, Rothman JK 1998. Surveillance in Modern Epidemiology. Maple Press Lippincott-Raven Publishers.

Cannon RM, Roe RT, 1982. Livestock Disease Surveys: A Field Manual for Veterinarians. Australian Government. Publishing service, Canberra.

Calba C, Ponsich A , Nam S, Collineau L, Min S, Thonnat J, Goutard F, 2014. Development of a participatory tool for the evaluation of Village Animal Health Workers in Cambodia. Acta Tropica. 134: 17–28.

Calba C, Goutard FL, Hoinville L, Hendrikx P, Lindberg A, Saegerman C, Peyre M, 2015. Surveillance systems evaluation: a systematic review of the existing approaches. BMC Public Health.15:448.

Calba C, Goutard FL, Vanholme L, Antoine-Moussiaux N, Hendrikx P, Saegerman C, 2016. The Added-Value of Using Participatory Approaches to Assess the Acceptability of Surveillance Systems: The Case of Bovine Tuberculosis in Belgium. PLoS One. 27;11(7):e0159041. doi: 10.1371/journal.pone.0159041. eCollection 2016.

Cameron AR, Baldock FC, 1998a. A new probability formula for surveys to substantiate freedom from disease. Prev Vet Med. 34(1):1-17.

Cameron AR, Baldock FC, 1998b. Two-stage sampling in surveys to substantiate freedom from disease. Prev Vet Med. 34(1):19-30.

Cameron AR, 2012. The consequence of risk-based surveillance: Developing output-based standards for surveillance to demonstrate freedom from disease. Prev Vet Med. 105: 280-286.

Claes G, Welby S, Van Den Berg T, Van Der Stede Y, Dewulf J, Lambrecht B, Marché S, 2013. The impact of viral tropism and housing conditions on the transmission of three H5/H7 low pathogenic avian influenza viruses in chickens. Epidemiol Infect. 141(11):2428-43. doi: 10.1017/S0950268813000125. Epub 2013 Feb 11

Comin A, Stegeman A, Marangon S, Klinkenberg D, 2012. Evaluating surveillance strategies for the early detection of low pathogenicity avian influenza infections. PLoS One. 7(4):e35956.

Cowie CE, Gortázar C, White PCL, Hutchings MR,Vicente J, 2015. Stakeholder opinions on the practicality of management interventions to control bovine tuberculosis. Vet J. 204:179–185.

de Koeijer AA, Boender GJ, Nodelijk G, Staubach C, Meroc E, Elbers AR, 2011. Quantitative analysis of transmission parameters for bluetongue virus serotype 8 in Western Europe in 2006. Vet Res. 24:42-53.

del Rocio Amezcua M, Pearl DL, Friendship RM, McNab WB, 2010. Evaluation of a veterinary-based syndromic surveillance system implemented for swine. Can J Vet Res. 74(4):2241-2451.

Drewe JA, Hoinville LJ, Cook AJ, Floyd T, Stärk KD, 2012. Evaluation of animal and public health surveillance systems: a systematic review. Epidemiol Infect. 140(4):575-590.

Drewe JA, Häsler B, Rushton J, Stärk KD, 2014. Assessing the expenditure distribution of animal health surveillance: the case of Great Britain. Vet Rec. 174(1):16.

Doherr MG, Audigé L, 2001. Monitoring and surveillance for rare health-related events: a review from the veterinary perspective. Philos Trans R Soc Lond B Biol Sci. 356(1411):1097-1106.

Dohoo I, Martin W, Stryhn H, 2012. Methods in Epidemiologic Research Hardcover. VER Inc. ISBN-10: 0919013732 ; ISBN-13: 978-0919013735

Dórea FC, McEwen BJ, McNab WB, Revie CW, Sanchez J, 2013.Syndromic surveillance using veterinary laboratory data: data pre-processing and algorithm performance evaluation. J R Soc Interface. 83:20130114.

Dufour B, Hendrikx P, 2007. Epidemiological surveillance in animal health. Ed. Cirad Aeema.

Elbers AR, Fabri TH, de Vries TS, de Wit JJ, Pijpers A, Koch G, 2004. The highly pathogenic avian influenza A (H7N7) virus epidemic in The Netherlands in 2003--lessons learned from the first five outbreaks. Avian Dis. 48: 691-705.

Elbers AR, Gorgievski-Duijvesteijn MJ, Zarafshani K, Koch G, 2010. To report or not to report: a psychosocial investigation aimed at improving early detection of avian influenza outbreaks. Rev Sci Tech. 29(3):435-449.

European Commission (EC), 1964. Council Directive 64/432/EEC of 26 June 1964 on animal health problems affecting intra-Community trade in bovine animals and swine. Official Journal of the European Union L121:1977–2012.

European Commission (EC), 1992. Council Directive 92/117/EEC concerning measures for protection against specified zoonoses and specified zoonotic agents in animals and products of animal origin in order to prevent outbreaks of food-borne infections and intoxications. Official Journal of the European Union L62:38-48.

European Commission (EC), 2000. Regulation (EC) No 1760/2000 establishing a system for the identification and registration of bovine animals and regarding the labelling of beef and beef products and repealing Council Regulation (EC) No 820/97. Official Journal of the European Union L204:1-10.

European Commission (EC), 2003a. Council Directive 2003/99/EC of the European parliament and of the council of 17 November 2003 on the monitoring of zoonosis and zoonotic agents, amending Council Decision 90/424/EEC and repealing Council Directive 92/117/EEC. Official Journal of the European Union L325:31-40.

European Commission (EC), 2003b. Regulation (EC) No 1082/2003 of 23 June 2003 laying down detailed rules for the implementation of Regulation (EC) No 1760/2000 of the European Parliament and of the Council as regards the minimum level of controls to be carried out in the framework of the system for the identification and registration of bovine animals. Official Journal Journal of the European Union L156:9 –12.

European Commission (EC), 2004. Commission Regulation (EC) No 911/2004 of 29 April 2004 implementing Regulation (EC) No 1760/2000 of the European Parliament and of the Council as regards eartags, passports and holding registers. Official Journal Journal of the European Union L163:65 –70.

European Commission (EC), 2005. Commission Regulation (EC) No 644/2005 of 27 April 2005 authorising a special identification system for bovine animals kept for cultural and historical purposes on approved premises as provided for in Regulation

(EC) No 1760/2000 of the European Parliament and of the Council. Official Journal Journal of the European Union L107:18–19.

European Commission (EC), 2010. Commission Decision 2010/367/EU on the implementation by Member States of surveillance programs for avian influenza in poultry and wild birds. Official Journal of the European Union L166: 22-32.

European Commission (EC), 2012. Commission Regulation (EU) No 456/2012 of 30 May 2012 amending Regulation (EC) No 1266/2007 on implementing rules for Council Directive 2000/75/EC as regards the control, monitoring, surveillance and restrictions on movements of certain animals of susceptible species in relation to bluetongue. Official Journal of the European Union L141:7-13.

European Union (EU). Regulation 2016/429 of the European Parliament and of the Council of 2016-03-09 on transmissible animal diseases and amending and repealing certain acts in the area of animal health (Animal Health Law), Official Journal of the European Union L84, 1-208.

European Union (EU), 2017. EU institutions and other bodies, Law Making. URL (Last update date: 02/01/2017): https://europa.eu/european-union/about-eu/institutions-bodies_en#law-making (Consulted date: 02/01/2017).

Ensoy C, Aerts M, Welby S, Van der Stede Y, Faes C, 2013. A dynamic spatio-temporal model to investigate the effect of cattle movements on the spread of bluetongue BTV-8 in Belgium. PLoS One. 8(11):e78591.

Ensoy C, Faes C, Welby S, Van der Stede Y, Aerts M, 2014. Exploring cattle movements in Belgium. Prev Vet Med. 116(1-2):89-101.

Faes C, Litière S, Méroc E, Van der Stede Y, Mintiens K, Aerts M, 2011. Estimating herd prevalence on the basis of aggregate testing of animals. JR Statist Soc. A 174:155–174.

Farnsworth ML, Hamilton-West C, Fitchett S, Newman SH, de La Rocque S, De Simone L, Lubroth J, Pinto J, 2010. Comparing national and global data collection systems for reporting, outbreaks of H5N1 HPAI. Prev Vet Med. 95(3-4):175-85.

Federal public services (FPS), Health, Food Chain Safety and Environment, 2017. Les compétences fédérales en matière de santé animale et végétale. URL (Last update date: no date) http://www.health.belgium.be/fr/animaux-et-vegetaux (Consulted date: 02/01/2017).

Food and Agriculture Organization (FAO), 2017. Challenges for Livestock Disease Surveillance. URL (Last update date: 25/07/2016): http://www.fao.org/AG/AGAInfo/resources/en/pubs_ah.html (Consulted date: 02/01/2017).

Frössling J, Wahlström H, Agren EC, Cameron A, Lindberg A, Sternberg Lewerin S, 2013. Surveillance system sensitivities and probability of freedom from Mycobacterium avium subsp. paratuberculosis infection in Swedish cattle. Prev Vet Med. 108(1):47-62.

Gilbert M, Mitchell A, Bourn D, Mawdsley J, Clifton-Hadley R, Wint W, 2005. Cattle movements and bovine tuberculosis in Great Britain. Nature. 26;435(7041):491-496.

Graat EAM, de Jong MCM, Frankena K, Franken P, 2001. Modelling the effect of surveillance programmes on spread of bovine erpesvirus 1 between certifieded cattle herds. Vet Microbiol. 79(3):193-208.

Gustafson LL, Gustafson DH, Antognoli MC, Remmenga MD, 2013. Integrating expert judgment in veterinary epidemiology: example guidance for disease freedom surveillance. Prev Vet Med. 109(1-2):1-9.

Halasa T, Boklund A, 2014. The Impact of Resources for Clinical Surveillance on the Control of a Hypothetical Foot-and-Mouth Disease Epidemic in Denmark. PLoS ONE 9(7): e102480

Häsler B, Howe KS, Di Labio E, Schwermer H, Stärk KD, 2012. Economic evaluation of the surveillance and intervention programme for bluetongue virus serotype 8 in Switzerland. Prev Vet Med. 103(2-3):93-111.

Hendrikx P, Gay E, Chazel M, Moutou F, Danan C, Richomme C, Boue F, Souillard R, Gauchard F, Dufour B, 2011. OASIS: an assessment tool of epidemiological surveillance systems in animal health and food safety. Epidemiol Infect. 139(10):1486-1496.

Hoinville LJ, Alban L, Drewe JA, Gibbens JC, Gustafson L, Häsler B, Saegerman C, Salman M, Stärk KD, 2013. Proposed terms and concepts for describing and evaluating animal-health surveillance systems. Prev Vet Med. 2013 112(1-2):1-12.

Humblet MF, Moyen JL, Bardoux P, Boschiroli ML, Saegerman C, 2011.The importance of awareness for veterinarians involved in cattle tuberculosis skin testing. Transbound Emerg Dis. 58(6):531-536.

Jordan D, McEwen SA, 1998. Herd-level test performance based on uncertain estimates of individual test performance, individual true prevalence and herd true prevalence. Prev Vet Med. 36(3):187-209.

Knight-Jones TJ, Hauser R, Matthes D, Stärk KD, 2010. Evaluation of effectiveness and efficiency of wild bird surveillance for avian influenza. Vet Res. 41(4):50.

Marceau A, Madouasse A, Lehébel A, van Schaik G, Veldhuis A, Van der Stede Y, Fourichon C, 2014. Can routinely recorded reproductive events be used as indicators of disease emergence in dairy cattle? An evaluation of 5 indicators during the

emergence of bluetongue virus in France in 2007 and 2008. J Dairy Sci. 97(10):6135-6150.

Martin PA, Cameron AR, Greiner M, 2007a.Demonstrating freedom from disease using multiple complex data sources 1: a new methodology based on scenario trees. Prev Vet Med. 79(2-4):71-97.

Martin PA, Cameron AR, Barfod K, Sergeant ES, Greiner M, 2007b. Demonstrating freedom from disease using multiple complex data sources 2: case study--classical swine fever in Denmark. Prev Vet Med. 79(2-4):98-115.

Méroc, E, Herr, C, Verheyden, B, Hooyberghs, J, Houdart, P, Raemaekers, M, Vandenbussche, F, De Clercq, K and Mintiens, K, 2009. Bluetongue in Belgium: Episode II. Transbound and Emerg Dis. 56: 39–48. doi:10.1111/j.1865-1682.2008.01063.x

Méroc E, Strubbe M, Vangroenweghe F, Czaplicki G, Vermeersch K, Hooyberghs J, Van der Stede Y, 2012. Evaluation of the Salmonella surveillance program in Belgian pig farms. Prev Vet Med. 105(4):309-14. doi: 10.1016/j.prevetmed.2012.03.006. Epub 2012 Apr 7.

Mollenberghs, 2009. Survey methods. Diepenbeeck University, Master in applied statistics courses notes.

More SJ, Radunz B, Glanville RJ, 2015. Lessons learned during the successful eradication of bovine tuberculosis from australia. Vet Rec 177:224-232 doi:10.1136/vr.103163

Perrin JB, Ducrot C, Vinard JL, Morignat E, Calavas D, Hendrikx P, 2012. Assessment of the utility of routinely collected cattle census and disposal data for syndromic surveillance.Prev Vet Med. 105(3):244-252.

Queenan K, Häsler B, Rushton J, 2016. A One Health approach to antimicrobial resistance surveillance: is there a business case for it?Int J Antimicrob Agents. 2016 Jul 25. pii: S0924-8579(16)30183-2. doi: 10.1016/j.ijantimicag.2016.06.014. [Epub ahead of print]

Rabinowitz PM, Galusha D, Vegso S, Michalove J, Rinne S, Scotch M, Kane M, 2012. Comparison of human and animal surveillance data for H5N1 influenza A in Egypt 2006-2011. PLoS ONE 7(9): e43851.

Racloz V, Griota C, Stärk KDC, 2006. Sentinel surveillance systems with special focus on vector-borne diseases. Animal Health Research Reviews 7 (1-2): 71-79.

Roelandt S, Van der Stede Y, D'hondt, B, Koenen, F, 2015. The Assessment of African Swine Fever Virus Risk to Belgium Early 2014, using the Quick and Semi quantitative Pandora Screening Protocol. Transboundary and Emerging Diseases. doi: 10.1111/tbed.12365.

Rutten N, Gonzales JL, Elbers AR, Velthuis AG, 2012. Cost analysis of various low pathogenic avian influenza surveillance systems in the Dutch egg layer sector. PLoS One. 7(4):e33930.

Salmon MD, 2003. Surveillance and Monitoring Systems in Animal Disease Surveillance and Survey Systems, Ed. MD Salman, 2003 Iowa States Press.

Salman MD, Stärk KD, Zepeda C, 2003. Quality assurance applied to animal disease surveillance systems. Rev Sci Tech. 22(2):689-696.

Stärk KD, Salman M, Tempelman Y, Kihm U, 2002. A review of approaches to quality assurance of veterinary systems for health-status certification. Prev Vet Med. 56(2):129-140.

Stärk KD, Häsler B, 2015. The value of information: Current challenges in surveillance implementation. Prev Vet Med. 122(1–2): 229–234.

Valarcher JF, Leforban Y, Rweyemamu M, Roeder PL, Gerbier G, Mackay DK, Sumption KJ, Paton DJ, Knowles NJ, 2008. Incursions of foot-and-mouth disease virus into Europe between 1985 and 2006. Transbound Emerg Dis. 55(1):14-34.

Velthuis AG, De Jong MC, De Bree J, 2007. Comparing methods to quantify experimental transmission of infectious agents. Math Biosci. 210(1):157-176.

Vergne T, Paul MC, Chaengprachak W, Durand B, Gilbert M, Dufour B, Roger F, Kasemsuwan S, Grosbois V, 2014. Zero-inflated models for identifying disease risk factors when case detection is imperfect: application to highly pathogenic avian influenza H5N1 in Thailand. Prev Vet Med. 114(1):28-36.

Welby S, van den Berg T, Marché S, Houdart P, Hooyberghs J, Mintiens K, 2010. Redesigning the serological surveillance programme for notifiable Avian Influenza in Belgian professional poultry holdings. Avian Dis. 54:597-605.

Welby S, Imberechts H, Riocreux F, Dierick K, Bertrand S, Wildemauwe C, Hooyberghs J, Van der Stede Y, 2011. Comparison of Salmonella Enteritidis phage types isolated from layers and humans in Belgium in 2005. Foodborne Pathog Dis. 8:929-934.

Welby S, Govaerts M, Vanholme L, Hooyberghs J, Mennens K, Maes L, Van Der Stede Y, 2012. Bovine tuberculosis surveillance alternatives in Belgium. Prev Vet Med. 106(2):152-161.

Welby S, Méroc E, Faes C, De Clercq K, Hooyberghs J, Mintiens K, Van der Stede Y, 2013. Bluetongue surveillance system in Belgium: a stochastic evaluation of its risk-based approach effectiveness. Prev Vet Med. 2013 112(1-2):48-57. doi: 10.1016/j.prevetmed.2013.07.005. Epub 2013 Aug 12.

Welby S, van Schaik G, Veldhuis A, Brouwer H, Peroz C, Berends I, Fourichon C, Wever P, van der Stede Y., 2016. Cost efficiency of passive active and syndromic surveillance components for early detection of emerging vector borne disease in cattle: Bluetongue as case study for Belgium, France and the Netherlands. Transbound Emerg Dis. ,2016. doi: 10.1111/tbed.12564. [Epub ahead of print]

World Organisation for Animal Health (OIE), 2017a. Terrestrial Animal Health Code: . URL (Last update date: 2017): http://www.oie.int/en/international-standard-setting/terrestrial-code/ (Consulted date: 02/01/2017).

World Organisation for Animal Health (OIE), 2017b. Manual of Diagnostic Tests and Vaccines for Terrestrial Animals. . URL (Last update date: 2017): http://www.oie.int/en/international-standard-setting/terrestrial-manual/ (Consulted date: 02/01/2017).

CHAPTER II: AIMS OF THE THESIS

In the last years standards have evolved from input based (when it was compulsory for each member state and national authority to use the same tests and follow the same sampling scheme or frequency of testing) to adopt output based (where each member state is free to choose its own surveillance strategy but where it must also demonstrate it achieved the required objective; i.e., disease-free situation). Local, national and international standards detail minimum requirements and provide flexible guidelines allowing enough freedom to encompass all country specific epidemiological situations when implementing and designing national surveillance programmes. Thus, over time, rules have evolved from an obligation of means (sample 40 animals per farm no matter the herd size and herd characteristics; e.g., closed herd, limited movements versus large herd frequent movements) towards an obligation of results (demonstrate that one can detect a minimal prevalence of 30% with a 95% confidence level). This major shift to “fit for purpose surveillance”, has given rise to more freedom for member states to design and tailor surveillance systems according to national needs, local risk factors and husbandry practices allowing for heterogeneous population and epidemiological situation. However, to maintain mutual trust between countries for trading, economic and political purposes, having similar approaches in surveillance eases comparison and auditing, thereby offering firm guarantees. Thus, these changes implied the need for tools to measure whether output standards were delivered (Cameron, 2012; Salman et al., 2003). Responsibility is left to the risk managers to opt for the best testing scheme (test, sample size and sampling framework) that will enable achievement of required standards. Although this approach to surveillance has the potential of higher flexibility, auditing the cost efficiency performance of such surveillance will be more complex.

The present thesis aims at defining a comprehensive and objective quantitative approach for evaluation of notifiable animal disease surveillance systems by testing their performance against defined purposes and objectives, according to needs and priorities expressed by stakeholders involved and compared with other countries.

More specifically, the output based method (scenario tree methods) were adapted for four important animal diseases (avian influenza, bovine brucellosis, bovine tuberculosis, bluetongue) in Belgium for which the surveillance programmes had to

be adapted accordingly. The aim was to explore the different possibilities offered by these tools together with the advantages and disadvantages.

Specific issues addressed in the following chapters are:

- How can we define risk based surveillance, risk cause versus risk consequence?
- How to handle perception of risk and varying objectives?
- Can combining components and historical information over time increase negative predictive value of surveillance?
- What are the assumptions limiting assessment of surveillance performance: design prevalence, diagnostic process, input/output validation?
- How to conciliate performance and cost of surveillance systems: comparing surveillance strategies across countries?

CHAPTER III: REDESIGNING THE SEROLOGICAL SURVEILLANCE PROGRAMME FOR NOTIFIABLE AVIAN INFLUENZA IN BELGIAN PROFESSIONAL POULTRY HOLDINGS

Avian Dis. 2010 Mar; 54(1 Suppl):597-605.

Redesigning the serological surveillance program for notifiable avian influenza in Belgian professional poultry holdings.

Welby S, van den Berg T, Marché S, Houdart P, Hooyberghs J, Mintiens K.

Abstract

This study was aimed at redesigning the Belgian active surveillance program for domestic birds in professional poultry holdings based on a risk analysis approach.

Stochastic quantitative analysis, combining all data sources, was run to obtain sensitivity estimates for the detection of an infected bird in the different risk groups identified.

An optimal number of holdings for each risk group were then estimated on the basis of the different sensitivities obtained. This study was a useful tool for decision makers to reallocate the total amount of samples to be taken in the next year(s) in Belgium, thus optimizing the field resources, and improving efficiency of disease surveillance such as required by the international standards.

INTRODUCTION

Avian Influenza (AI) is caused by type A Influenza viruses, which are RNA viruses that belong to the Orthomyxoviridae family. Among type A, the viruses can be further sub-typed based on the characteristics of two viral proteins, i.e. hemagglutinin (HA) and neuraminidase (NA). Among all possible combinations between the 16 different haemaglutinins and 9 neuraminidases, to date only H5 and H7 have the potential of causing highly pathogenic avian influenza (HPAI). Any highly pathogenic avian influenza isolate is classified as notifiable avian influenza (NAI) virus. Although all virulent strains isolated up to now have been either of the H5 or H7 subtype, most H5 or H7 isolates have been of low virulence. Due to the risk of a low virulent H5 or H7 becoming virulent by mutation in poultry hosts, all H5 and H7 viruses have also been classified as NAI viruses (OIE, 2008). Wild birds are a natural reservoir of LPAI and can introduce it into domestic poultry populations, especially in outdoor birds (Hinshaw et al., 1980). Moreover, wild birds can transport the HPAI across borders when migrating and spread the infection in different countries along their migratory route (Alexander., 2000, 2007a, b; Gilbert et al., 2006a, b). As a secondary route of spread, human behavior is probably the main reason responsible for the maintenance of the infection within certain areas, through intensive farming and inappropriate bio-security measures, as well as trading and smuggling (Alexander, 2007a,b; Elbers et al., 2008; Gilbert et al., 2006a,b; van den Berg and Houdart, 2008). A worldwide pandemic of HPAI H5N1 occurred in 2003-2004, which had its

origin in a low-level but endemic circulation of the virus in Chinese ducks which started in 1996. Following the devastating economical and ethical repercussions the eradication of HPAI had in Italy, it was suggested that the European Member States must set up survey programs in order to detect the LPAI before it actually spreads from farm to farm causing sanitary problems and carrying the risk of mutation into HPAI (Capua and Marangon, 2000). Also the outbreak of AI in the Netherlands led to the conclusion that serological monitoring of LPAI was needed in addition of reporting of HPAI (EC, 2005). Similarly in Belgium, it was suggested that adaptive measures to detect the virus before it spreads must be set in order to prevent the use of drastic measures (van den Berg and Houdart, 2008). In addition, the fear of a huge pandemic in humans and animals together with its dramatic consequences in the poultry industry underlined the need of detecting the pathogen before it actually spreads. Decision 2007/268 of the European Commission (EC, 2007) modifying the decision 2004/450/CE prescribes compulsory surveillance programs to be implemented in all European Member States for the detection of AI. This decision insists as well on targeting the surveillance in particular populations more at risk.

The Belgian surveillance program for AI is carried out in accordance with the Commission Decisions 2005/734 and 2007/268 (EC 2005, 2007). It concerns both domestic poultry (e.g. chickens, turkeys, geese, ducks, ratites, other poultry (guinea fowl, partridges, pheasants, meet pigeons)) and wild birds and consists in both passive and active surveillance. Commission Decision 2005/734 (EC, 2005) aims at an early detection system in professional poultry holdings to detect an introduction of an infection with HPAI serotypes H5N1 or H7N7. Commission Decision 2007/268 (EC 2007) sets up guidelines for active surveillance in professional poultry holdings: 1) detection of sub-clinical infections of LPAI H5, H7 2) targeting of the detection of LPAI H5, H7, in poultry populations which are more at risk due to their exposure, their rearing method, or due to their specie susceptibility 3) substantiation of freedom of disease status from a particular region or zone, in the context of international trade. Commission Decision 2007/268 (EC, 2007) also implements the standard guidelines for implementing passive and active surveillance in wild birds. The active surveillance in domesticated poultry aims at detecting the LPAI circulation through serological sampling except in ratites where the sample collection is cloacal and tracheal swabbing. The active surveillance in wild birds consists in cloacal and

tracheal swabbing during the ringing of wild birds and during the hunting season. The passive surveillance aims at early detecting HPAI in professional poultry holdings and is based on the reporting of any increase in mortality or morbidity or any decrease in water consumption, feed intake, or egg production. In wild birds the passive surveillance consists in the reporting of any suspicious increase of mortality in wild birds to the Royal Belgian Institute of Natural Sciences (RBINS).

The risk for a bird population to become infected with AI is influenced by numerous factors. The presence of migrating wild birds (usually characterised by presence of larger water surfaces), high population density, outdoor housing, regions with bird movements, numerous external contacts with other farms, inappropriate bio-security measures may all increase the risk for infection. The response to infection also differs between birds' species. Infected chickens and turkeys will most often show clear clinical symptoms whereas geese and ducks will have more sub-clinical infections. The fact that virus can circulate in the latter species without notice and thus can be transmitted to poultry where mutation can occur into a HPAI, makes them high risk species (Alexander 2000; Bavinck et al., 2009; Brown et al., 2003; Capua et Marongon, 2000; Elbers et al., 2004; Fang et al., 2008; Homme et al., 1970; Tumpey et al., 2004). All these factors which influence the probability of infection should be considered in the design of a risk based surveillance system in a country, as the sampling should focus on sub-populations with higher probability of infection.

The aim of this study was to evaluate the sensitivity of the Belgian serological active surveillance program for detecting LPAI in professional poultry holdings. In other words, to test the capability of the current active serological surveillance program to detect the infection if the country is infected. The sensitivity was estimated for different risk groups (sub-populations): based on region (five different regions were considered according to the relative risk zone surface within each region), risk zone (defined as zones where in between farm distance and likelihood of disease introduction was higher), holdings with or without outdoor facilities and species. Secondly, we aimed to propose targeted surveillance protocol to allow an optimal probability of detecting an infection in the different risk groups. This optimal sample size took into account the sensitivities that were obtained for the current surveillance and the legal requirements. An optimised sample scheme, redistributing the samples foreseen to be taken in 2009, was proposed following the optimal results obtained.

Martin et al (2007) have proposed methods for the analysis of complex surveillance systems which are in place for substantiating freedom from disease. These studies have been led in the past for different diseases and proved their efficiency (Hadorn et al., 2008; Martin et al., 2007a,b; Stark et al., 2006). The consecutive events which contribute to the detection of an infection in a population are described as nodes in a scenario tree. Each node is related to a risk for infection or probability of detection and following their differential risk of infection or probability of detection, respective reference population proportions (PPr) and the number of animals processed through the surveillance system component (SPr) are attributed to each category node in the tree. PPr defines the population proportion within each branch to the overall reference population in that category node, whilst SPr defines the number of sampled population in that branch. Data out of passed experience, literature figures as well as expert elicitation can be used to provide the input parameters in the tree.

MATERIAL AND METHODS

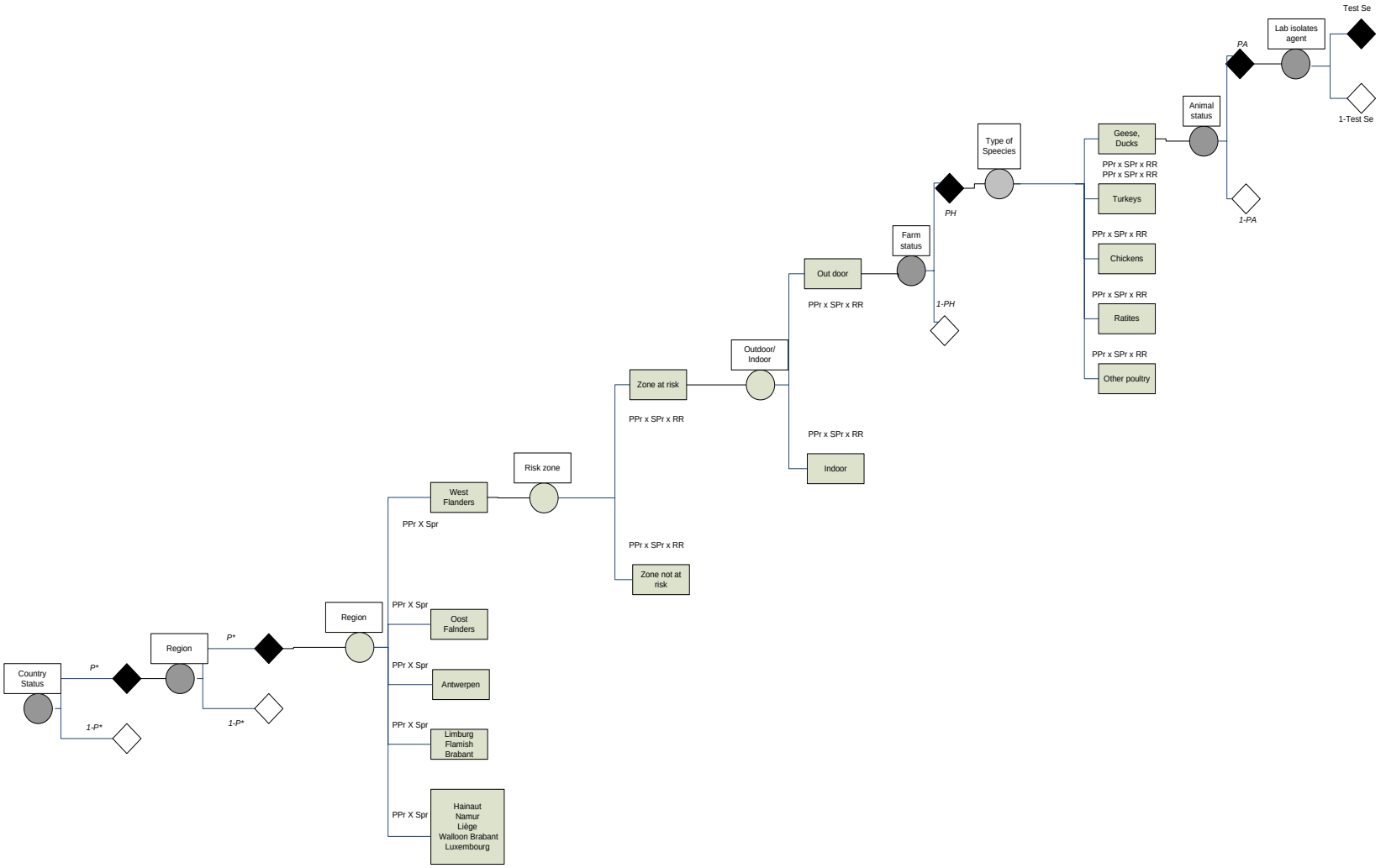
Active surveillance for LPAI in domesticated poultry

Domesticated poultry over 13 weeks of age and present on professional holdings are considered as study population. The sampling frame consists of Belgian professional holdings as identified in the Animal Identification and Registration System (SANITEL) with more then 200 birds or with more than 15 ratites, excluding holdings with broilers. The sampling is stratified by regions (FASFC, 2007) to ensure representative sampling of the total population. The number of holdings to sample per region must be sufficient to ensure the detection of an infected holding with 95% confidence, if 5% (design prevalence) of the holdings are infected, except for turkey, geese and ducks where a 99 % confidence level is required. The number of birds to be sampled per holding must guarantee the detection of a 30% design prevalence with a 95% confidence level for all poultry types, except for turkey, geese and ducks where a fixed number of 40-50 animals must be sampled in each holding. These requirements were set by the Community Reference Laboratory (VLA, Weybridge, UK) for the AI Monitoring program. All holdings are sampled once per year, except holdings in risk zones or with turkeys, geese and ducks at risk which are sampled twice per year.

Scenario tree description

The methods proposed by Martin et al (2007) were adopted for estimating the probability of detecting LPAI in case it would be introduced in one of the risk populations in Belgium. A scenario tree was designed in Microsoft Visio comprising nodes, branches, and outcomes (Figure 1). Three different types of nodes were used: 1) category nodes distinguish different population proportions homogeneous with regard to the risk of infection or detection; 2) infection nodes involve steps in the process related to the introduction of infection; 3) detection nodes involve steps in the process of detecting the infection. They characterise the different effective probability of infection or detection (i.e.; farm status, animal status and lab sensitivity). The tree starts with country status as infection node, then region status also an infection node, followed by a category node with five branches representing the regions that were established for the sampling stratification in Belgium. Then category nodes 'risk zone' and 'outdoor facilities', with each two branches, classify the different holdings according to the fact of being in a risk zone or having outdoor facilities, respectively. As following, the farm status infection node indicates the effective probability for a farm to be infected, within each branch. From farm status, a category node 'species' differentiates bird's species within farms according to the risk of infection represented by the different species. At the end of each limb of the tree, different risk groups according to risk zone, outdoor facilities, and type of species are obtained, each having a specific effective probability of infection and a probability of detection. A risk group is defined by the set of possible combination of the different categories. For instance a risk group RiskZoneOutdoorChicken combines the categories chicken in category farm with outdoor facilities in category risk zone. Separate trees were developed for each region in Belgium and for the whole country.

Figure 1: Scenario tree representing the different steps between the introduction of LPAI in domestic poultry in Belgium and it's detection by the active surveillance system



Parameters in the different nodes

In order to estimate the sensitivity of detecting an infection in the different risk groups, parameters were allocated to the different categories in the tree. Population proportions (PPr) of holdings or animals and the number of holdings or animals sampled (SPr) were allocated to each category node or the different risk groups respectively. Population proportions (PPr) of holdings or animals and the holdings or number of animals processed by the surveillance component (SPr) were obtained from the SANITEL dataset.

To obtain the effective probability of infection for holdings (EPIH) and animals (EPIA) in the different risk groups, the overall herd (PH) and within-herd (PA) design prevalence as imposed by Commission Decision 2007/268 (PH=5 % and PA=30 %) was weighted by a risk group specific (adjusted) relative risk. The relative risks for infection regarding being in a risk zone, having outdoor facilities, species at risk, lab sensitivity were obtained from either passed experience or literature. The AR's (equation 1) were obtained by dividing the relative risks (RR) for infection by the sum of the relative risk in each risk group multiplied by the PPr in each branch of the category.

$$AR_i = \frac{RR_i}{\sum_{i=1}^{n} RR_i * PPr_i} \quad (1)$$

For each risk group i, the effective probability of infection for holdings (EPIH) and animals (EPIA) was calculated as follows (equation 3, 4):

$$EPIH = AR_i * PH \quad (2)$$

$$EPIA = AR_i * PA \quad (3)$$

Per risk group, the sensitivity of detecting an infected holding (equation 4) and an infected animal within a holding (equation 5) was calculated taking the EPIH and EPIA, the sensitivity of the laboratory diagnosis (Lab Se), the number of animals sampled (na), and the number of holdings sampled (nh) into account.

$$Holding\ Se_i = 1 - (1 - (EPIA_i * Lab\ Se))^{n_a} \quad (4)$$

$$Risk\ Group\ Se_i = 1 - (1 - (EPIH_i * Holding\ Se_i))^{n_a} \quad (5)$$

The Lab Se was derived from the test Se, obtained from a study done in the Netherlands on the inhibition of haemagglutination (Koch, 2003). The specificity wasn't included as it assumed to be 1, which can be motivated by the fact that each positive result will be further investigated until it truly shows positive.

Rationale and methods for obtaining sensitivities

The scenario trees were modelled in Microsoft Excel. Separate trees were constructed in different spreadsheets for each region and for the whole of Belgium. Within each spreadsheet, different worksheets were created to facilitate the modelling exercise. All the inputs of the population proportion, relative risk, and design prevalence were introduced in an 'Input/Output' worksheet to calculate the adjusted risks, effective probabilities of infection and detection (equations 1 and 2). In a 'Proportion Calculation' worksheet, the actual SANITEL dataset was worked through in order to obtain the different populations (representative and sampled) counts in different categories with the help of pivot tables. In a 'Tree' worksheet the effective probabilities of infection obtained in the Input/Output sheet was recalled to calculate the sensitivities for the different risk groups taking the sampling fraction in the different risk groups into account (equations 3 and 4).

The outputs were computed using @risk software (Palisade, 2007) taking the uncertainty and variability of parameters into account by fitting appropriate parameter distributions. The sensitivity estimates for the different risk group were obtained by separate hypergeometric simulation for each region with 10000 iterations in each simulation. This offers the opportunity to consider all the possible pathways in the scenario by sampling from the parameter distributions.

Rationale and methods for sample size calculations

An optimal sample size was calculated per risk group using the methods proposed by Cameron and Baldock (1998) using the Freecalc software (Survey Tool box (Ausvet, 2008)) based on the effective probabilities of infection and the sensitivities that were obtained from the scenario trees. The optimal sample sizes were estimated for both the number of samples within holdings as well as the number of holdings.

Here again the specificity was assumed to be 100%. For the Type I and Type II error (for falsely rejecting or accepting the null hypothesis of presence of infection) the desired confidence levels in Commission Decision 2007/268 (EC, 2007) were used, i.e. 95% for all bird species except turkeys, geese and ducks where a 99% confidence level is required.

Because in 2010, Belgium had an EU provision for taking 10000 samples for avian influenza surveillance, based on the estimated optimal sample size, we proportionally redistributed the 10000 samples accordingly (optimised sample size). This

proportional redistribution over the different risk groups enabled to achieve a risk-based surveillance program.

RESULTS

Descriptive statistics study population

Population proportions (PPr) and sampled populations proportions were obtained from the Sanitel dataset. A total of 47,395,827 birds present on 2,097 professional holdings were considered in this study and 13,450 birds in 2,016 holdings were sampled. Population proportions were allocated to category nodes, whilst numbers of animals processed were mainly attributed to risk group nodes which enabled the sensitivity calculation of the given risk group. Table 1 shows population total and relative proportions to each category node in Belgium whereas Table 2 shows the number of birds and holdings sampled in each risk group. Holdings were allocated to risk groups with higher risk if a mixture of risk groups were present on the holding, e.g. if a holding had a minor number of birds reared in outdoor facilities next to a majority of birds raised inside then it was classified as holding with outdoor facilities.

Table 1: Proportion of total domesticated bird population in each category

Risk category	***Population size***	***Proportion of total Domesticated bird Population***
Risk Zone Population	3,589,154	0.076
Non Risk zone population	43,806,673	0.924
Outdoor facilities population	9,252,265	0.195
Indoor facilities population	38,143,562	0.805
Chicken	46,355,175	0.978
Turkey	317,371	0.007
Geese and Ducks	79,628	0.002
Ratites	0	0
Other Poultry	643,653	0.013

Table 2 Number of birds and holdings sampled in each risk group

Category	# holdings sampled in Risk zone	# birds sampled per holding in Risk Zone	# holdings sampled in non Risk zone	# birds sampled per holding in non Risk zone
Outdoor Chickens	6	10x2*	140	10
Outdoor Turkeys	2	10x2*	9	10x2*
Outdoor Geese and Ducks	0	50x2*	8	50x2*
Outdoor Ratites	2	10x2*	38	10
Outdoor Other Poultry	0	10x2*	16	10
Indoor Chickens	46	10x2*	564	10
Indoor Turkeys	4	10x2*	36	10x2*
Indoor Geese and Ducks	0	50x2*	26	50x2*
Indoor Ratites	0	10x2*	20	10
Indoor Other Poultry	4	10x2*	30	10

*Twice yearly

Input parameters

For estimating the relative risk of being in a risk zone compared to a non-risk zone, the between-farm distance was estimated to be influential. Therefore the spatial transmission characteristics (representing the different transmission ratios according to the in between farm distance) of the 2003 AI epidemic in the Netherlands as presented by Boender et al.(2007), were used to obtain the RR for the risk zone compared to the non-risk zone. The choice of appropriate distributions was estimated following Vose (2000). A triangular distribution TR(0.46; 12; 16) was used to express the RR in the risk zone whereas a uniform distribution U(1, 1) was used for the RR in the non-risk zone. The RR for having outdoor facilities was derived from data presented by Elbers et al. (2004). In this study the evolution of the 2003 Dutch AI

epidemic was investigated in 5 different farms and the housing system was found to be an explanatory factor for differences in mortality rate between farms. Out of the different mortality rates, a RR was approximated with this explanatory variable and was introduced as a triangular distribution TR (0.1; 1; 5) in our model. Again a uniform distribution U (1,1) was used for the RR in farms with no outdoor facilities. The RR associated with the different birds species were derived from proportional mortality rates in the different species holdings of the EU Member States that were affected by the 2003 AI epidemic (Brown, 2003). Chickens were considered as reference group and therefore the RR was expressed as an uniform distribution U (1; 1). For the other species the RR's were again expressed as triangular distributions:

- Turkeys: TR (0.3-0.7-8);
- Geese and ducks: TR (1; 1.6; 7);
- Ratites: TR (0.8; 1.1; 1.4);
- Other poultry: TR (1.1; 1.5; 3.1).

The resulting AR's and EPI's for all risk groups are summarized in Table 3.

Table 3: Relative risk, adjusted risk and effective probabilities of infection

Risk category	***Relative Risk***	***Adjusted Risk***	***Effective probabilities of infection***
Holding in Risk Zone with Outdoor facilities	RiskTR (0.46, 12, 16)	5.775	0.49
Holding in Risk Zone with Indoor facilities	RiskU (1,1)	0.609	0.24
Holding in Non Risk with Outdoor facilities	RiskTR (0.1, 1, 5)	1.692	0.05
Holding in Non Risk with Indoor facilities	RiskU (1, 1)	0.832	0.02
Chicken	RiskU (1, 1)	0.972	0.29
Turkey	RiskTR (0.3, 0.7, 8)	2.915	0.87
Geese and Ducks	RiskTR (1, 1.6, 7)	3.109	0.93
Ratites	RiskTR (0.8, 1.1, 1.4)	1.069	0.32
Other Poultry	RiskTR (1.1, 1.5, 3.1)	1.846	0.55

The input parameter for Lab sensitivity of the inhibition of haemagglutination was expressed as a beta distribution beta (0.46; 0.03), based on the sensitivity with 95% confidence interval that was estimated by Koch (2003).

Estimated category effective probabilities of infection, sensitivities, and sample sizes per risk group.

In Table 3 as mentioned above the results of the different relative risk, adjusted risk and effective probabilities of infection are shown. The initial sample sizes, the optimal

sample sizes obtained following our model and the risk-based redistribution of the 10,000 samples foreseen for 2009 in Belgium (optimised sample size) are stated in Table 4. We notice that the effective probability of infection is higher in risk zones with outdoor facilities, and the species more at risk is geese and ducks.

Table 4: Summary for Belgium of the sensitivities, initial as well as optimal and optimised sample sizes (SS) estimation obtained in each risk group in Belgium

RG	Risk zone				Non risk zone			
	Se	Initial SS	Optimal SS	Optimised SS	Se	Initial SS	Optimal SS	Optimised SS
Out Chicken	0.99	120	70	450	0.99	1400	90	580
Out Turkey	0.93	40	30	200	0.61	180	60	380
Out GD	0	0	0	0	0.57	800	300	2000
Out Ratites	0.93	40	50*	250*	0.86	380	90	580
Out OtherP	0	0	0	0	0.57	160	80	500
In Chicken	0.99	920	90	580	0.99	5640	90	580
In Turkey	0.89	80	30	200	0.84	720	40	260
In GD	0	0	0	0	0.74	2600	200	1300
In Ratites	0	0	0	0	0.39	200	200	1300
In OtherP	0.88	80	40	260	0.54	300	90	580

The tables for the different regions can be found in the appendix. From Table 4 it can be observed that under sampling occurred in risk zones and species at risk but over sampling occurred in non-risk zones.

Table 5 is a summary of the actual optimised total sample size in each category group redistributed according to the number of holdings, and birds per holding to be sampled.

Table 5: Summary for Belgium of the number of holdings to sample in each risk group and the number of samples/holding in each risk group

RG	Risk zone		Non risk zone	
	Holdings SS	Animal/Holding SS	Holdings SS	Animal/Holding SS
Out Chicken	45	10	58	10
Out Turkey	20	10	38	10
Out GD	0	50	40	50
Out Ratites	25	10	58	10
Out OtherP	0	10	50	10
In Chicken	58	10	58	10
In Turkey	20	10	26	10
In GD	0	50	130	10
In Ratites	0	10	130	50
In OtherP	26	10	58	10

DISCUSSION

Modelling the active surveillance of LPAI in domestic birds by using scenario trees is a useful tool to estimate the sensitivity of detection in the different risk groups (Hadorn et al., 2008; Martin et al., 2007a,b; Stark et al., 2006). Based on the risk-group specific EPI's and Se's that resulted from the model a risk-based reallocation of the current sampling protocol was possible. This increases the confidence in detecting a probable new infection with the same total number of samples to be taken. This corroborated what was expected. In certain risk groups the estimated

ideal sampling size decreased drastically such as in chicken farms without outdoor facilities in non-risk zones. This can be easily understood, as in those highly density populated holdings, if the virus was present; it would spread fairly quickly, and therefore enhance the chance of detection. In other risk groups the initial sample size was underestimated such as chickens raised in risk zones with outdoor facilities. Thus it is important to increase the sampling in those populations. The larger sample size required in ratites reflects the low probability of infection in that population together with the very low initial sample size in those populations. This must be interpreted cautiously. Because those populations are so small, it might actually be totally insignificant to sample in those categories, as they would probably not constitute a high risk.

The Commission Decision 2007/268 (EC, 2007) insists on having an active surveillance system in domesticated birds to substantiate freedom of disease, as well as an early detection system and prevalence estimation of circulating LPAI. The sampling design imposed by this Decision is only able to substantiate freedom of disease with a certain confidence level. Therefore this study aimed at evaluating the efficacy of the surveillance system in substantiating freedom of disease, with a statistical confidence level based on risk quantification. The sampling design disregards sampling in time, and the likelihood of disease introduction between different sampling periods. The actual surveillance system substantiates freedom from H5 or H7 LPAI strains with yearly intervals. Using Bayes theorem, the results of the surveillance programme of previous years can be used to enhance its confidence during the following years (Martin et al., 2007a,b). Also accounting for disease introduction from year to year would be interesting to consider in a future study.

For an early detection system of circulating LPAI, a different sampling design is required. First, sampling would need to be repeated frequently, taking the time between incubation and seroconversion for LPAI into account. Moreover, sampling would need to be intensified or could even be restricted during periods or in subpopulations with higher risk for introduction of infection. For example sampling could be intensified during migration season for wild birds, within migration zone (risk zones), in densely populated areas, in holdings with outdoor facilities or in areas where more poultry transport occurs. The results of this study provide insight on risk groups where intensifying the sampling would be appropriate. In Belgium

concentrating the sampling from mid-September to end of October and from the beginning of March to the end of April (Moniteur Belge, 2008) in risk zones and more specifically in holdings with outdoor facilities, and in species at risk such as turkeys, geese and ducks, would constitute an optimal early detection system. Also intensified sampling during cold winters could be suggested as those periods are considered at higher risk as migrating geese from Siberia and Denmark tend to come down south. Those birds are susceptible to carry AI. The proposed methodology can be extended to evaluate the efficiency of the early detection system by taking the repeating sampling and the repeating probability of introduction into account, of ongoing surveillance results. This would require a dynamic approach. Alternatively, a node season could be incorporated in the static approach, thus representing the sampling during migration periods, which could constitute a way to account for early detection.

For early detection of HPAI passive surveillance based on clinical findings is more relevant as stipulated in the Commission Decision 2005/734 (EC, 2005).

The objective of Commission Decision 2007/268 (EC, 2007) is also to estimate the design prevalence in order to estimate a target prevalence to reach in each Member State. When estimating prevalence in a population, diagnostic test sensitivity and specificity are two crucial parameters. The set guidelines of the commission is mainly aimed at detecting the infection if were present, thus focusing on the minimum sample required for detecting the disease if it was present. In order to estimate prevalence of disease the sample size would be different (Dohoo 2003; Thrushfield, 1995).

All these figures must be cautiously interpreted, as the model was built on assumptions which were not scaled. The distributions for the input parameters were obtained from available information and not from empirical data. Sometimes this might have biased the imputations. Laboratory sensitivity was derived from studies aiming at evaluating detection of antibody directed against HPAI influenza and not LPAI for which the sensitivity might be very different (Koch, 2003). Similarly the estimation of transmission ratio was based on spatio-temporal characteristics of the HPAI epidemic in the Netherlands, but those characteristics were mainly based on farm density and not really on disease introduction or migratory pathways, and does not reflect the strengthened biosecurity measures in place following these epidemics (Boender et al., 2007). Also the housing systems characterising the farm type node,

was estimated from difference observed between five farms with different housing systems (Koch and Elbers, 2006). This provided good estimation of disparities in disease evolution within different holdings but does not truly reflect the pattern of disease for LPAI evolution specifically in farm type as defined in the Tree. Finally species at risk were estimated from proportions of infected holding in each Member State. These differences were assumed to be linked to species sensitivity, but other hidden factors could have potentially influenced those results. The housing systems and contact pattern for the different species holdings might have been very different, and thus influenced this outcome. Nevertheless all these inaccurate estimates remain valuable data. Taking into account this uncertainty around our estimates, by fitting appropriate distributions to represent the range of possible outcomes, reinforces the value of this data. This still remains the most objective data to estimate the relative risk from one species to another. In the future, studies purposely designed to obtain those figures would strengthen the power of such model.

Taking the risk of introduction into account, different pathways (import, wild birds, migration period) should be investigated (Elbers et al., 2004; Fang et al., 2008; Gilbert et al., 2006b; Webster et al., 2006). Emphasizing the surveillance programs to LPAI backyard poultry and the proper application of biosecurity measures is important to consider too (Bavinck et al., 2009; de Wit et al., 2004; Capua and Marongon, 2000). Human behaviour constitutes the secondary route of spread, therefore respecting proper biosecurity measures is important (Bavinck et al., 2009, Bos et al., 2009, Elbers et al., 2004). The fear of repressive measures or negative consequences can bring farmers more reluctant to report or accept sampling of their poultry. The relationship between the farmer and the authorities as well as with the official vets is important to consider, too. It is important to integrate this aspect in the surveillance program as this will further enhance the sensitivity evaluation and therefore reorient efforts in a more efficient surveillance program (Elbers et al., 2008). Also an important issue is how confident we are in substantiating disease freedom with the set requirements by the EU Decision 2007/268 (EC, 2007). As the sampling scheme proposed by the Decision is directed to ensure the identification of an infection if 5% of the holdings are infected which in Belgium would mean 105 holdings. Because decreasing these design prevalence's would involve further costs, targeting in risk group where disease is more likely to cluster due to biological or

farming system reasons would be an alternative to enhance the efficacy of detection whilst keeping the cost and field work at a feasible level.

This study has enabled the identification of risk groups where disease tends to cluster according to biological, ecological or housing systems differences. Even though data used for the quantification of these differential risk was not based on empirical data, the fact of accounting for this inaccuracy by fitting appropriate distributions and taking in account this uncertainty, actually enhances the value of the risk estimate in a first stage and sensitivity in a later stage. Being able to quantify this sensitivity accounting for the complexity of the surveillance system provides great advantage to such models. Furthermore the more accurate sensitivity obtained through this model enables us then to conduct a more efficient target sampling. For assessing surveillance system in a country, this model is a valuable tool as it enables not only to provide a sensitivity but also a distribution around this sensitivity, thus substantiating freedom from disease with set statistical confidence level as required by the international standard guidelines.

Acknowledgements

We wish to thank Tony Martin, from the Australian Biosecurity Cooporative Research Centre, for his continuous support to the development of the model.

REFERENCES

Alexander DJ., 2000. A review of avian influenza in different bird species. Vet Microbiol 74: 3-13.

Alexander DJ., 2007a. An overview of the epidemiology of avian influenza. Vaccine 25: 5637-5644.

Alexander DJ, 2007b. Summary of avian influenza activity in Europe, Asia, Africa, and Australasia, 2002-2006. Avian Dis. 51: 161-166.

AUSVET (Australian Animal Health Service). Survey Tool Box: FreeCalc. URL (last update date: no date): http://www.ausvet.com.au/content.php?page=res_software#st, (consulted date: 05/03/2008).

Bavinck V, Bouma A, van Boven M, Bos ME, Stassen E, Stegeman JA, 2009. The role of backyard poultry flocks in the epidemic of highly pathogenic avian influenza virus (H7N7) in the Netherlands in 2003. Prev Vet Med. 88: 247-254.

Boender, GJ, Hagenaars TJ, Bouma A, Nodelijk G, Elbers AR, de Jong MC, van Boven M, 2007. Risk maps for the spread of highly pathogenic avian influenza in poultry. PLoS Comput Biol. 3: e71.

Bos ME, Nielen M, Koch G, Bouma A, De Jong MC, Stegeman A, 2009. Back-calculation method shows that within-flock transmission of highly pathogenic avian influenza (H7N7) virus in the Netherlands is not influenced by housing risk factors. Prev Vet Med. 88: 278-285.

Brown I, 2003. Surveys for avian influenza in poultry and wild birds in member states 2003. In: Proceedings of the Joint 9th Annual Meetings of the National Laboratories for Newcastle Disease and Avian Influenza of EU Member States. Alexander D.J., eds.: Brussels.

Capua I, Marangon S, 2000. The avian influenza epidemic in Italy, 1999-2000: a review. Avian Pathol. 29: 289-294.

de Wit JJ, Koch G, Fabri TH, Elbers AR, 2004. A cross-sectional serological survey of the Dutch commercial poultry population for the presence of low pathogenic avian influenza virus infections. Avian Pathol. 33: 565-570.

Dohoo I, Martin W, Stryn H, 2003. Veterinary Epidemiologic Research. AVC Inc.: Charlottetown, Canada. 34-36: 55-56.

European Commission (European Commission), 2007. Commission Decision on the implementation of surveillance programmes for avian influenza in poultry and wild birds to be carried out in the Member States and amending Decision 2004/450/EC (2007/268/EC). In: Official Journal of the European Union.

European Commission (European Commission), 2005. Commission decision laying down biosecurity measures to reduce the risk of transmission of highly pathogenic avian influenza caused by Influenza virus A subtype H5N1 from birds living in the wild to poultry and other captive birds and providing for an early detection system in areas at particular risk (2005/734/EC). In: Official Journal of European Union.

Elbers AR, Fabri TH, de Vries TS, de Wit JJ, Pijpers A, Koch G, 2004. The highly pathogenic avian influenza A (H7N7) virus epidemic in The Netherlands in 2003--lessons learned from the first five outbreaks. Avian Dis. 48: 691-705.

Elbers AR, Gorgievski M, van der Velden P, 2008. To report or not to report: opinions and attitudes of Dutch pig farmers and veterinary practitioners regarding clinical situations possibly caused by a notifiable pig disease. In: Proceedings of the Epizone 2nd annual meeting “Need for Speed”. Brescia (Italy).

Fang LQ, de Vlas SJ, Liang S, Looman CW, Gong P, Xu B, Yan L, Yang H, Richardus JH, Cao WC, 2008. Environmental factors contributing to the spread of H5N1 avian influenza in mainland China. PLoS ONE 3: e2268.

Federal Agency for the Safety of the Food Chain (FASFC), 2007. Monitoring program of Avian Influenza for 2007.

Gilbert M, Chaitaweesub P, Parakamawongsa T, Premashthira S, Tiensin T, Kalpravidh W, Wagner H, Slingenbergh J, 2006a. Free-grazing ducks and highly pathogenic avian influenza, Thailand. Emerg Infect Dis. 12: 227-234.

Gilbert M, Xiao X, Domenech J, Lubroth J, Martin V, Slingenbergh J, 2006b. Anatidae migration in the western Palearctic and spread of highly pathogenic avian influenza H5NI virus. Emerg. Infect. Dis. 12: 1650-1656.

Hadorn DC, Stark KD, 2008. Evaluation and optimization of surveillance systems for rare and emerging infectious diseases. Vet Res 39: 57.

Hinshaw VS, Webster RG, Turner B, 1980. The perpetuation of orthomyxoviruses and paramyxoviruses in Canadian waterfowl. Can. J. Microbiol. 26: 622-629.

Homme PJ, Easterday BC, Anderson DP, 1970. Avian influenza virus infections. II. Experimental epizootiology of influenza A-turkey-Wisconsin-1966 virus in turkeys. Avian Dis. 14: 240-247.

Koch G, Elbers ARW, 2006. Outdoor ranging of poultry: a major risk factor for the introduction and development of High-Pathogenicity Avian Influenza. NJAS journal. 54: 179-194.

Koch G, 2003. AI test validation. In: Proceedings of the 9th joint annual meetings of the National Reference Laboratories for Newcastle Disease and Avian Influenza. Alexander D.J., eds.: Brussels.

Martin PA, Cameron AR, Greiner M, 2007a. Demonstrating freedom from disease using multiple complex data sources 1: a new methodology based on scenario trees. Prev Vet Med. 79: 71-97.

Martin PA, Cameron AR, Barfod K, Sergeant ES, Greiner M, 2007b. Demonstrating freedom from disease using multiple complex data sources 2: case study--classical swine fever in Denmark. Prev Vet Med. 79: 98-115.

Moniteur belge, 2008. Avis relatif à la fixation de la période de migration des oiseaux migrateurs.- Exécution de l'arrêté ministériel du 03 avril 2006 portant sur les mesures temporaires de lutte contre l'influenza aviaire. Published the 28/04/2008; 29/02/208; 09/11/2007; 24/09/2007.

Palisade, 2007, The world's most popular risk analysis tool: @risk 4,5 version.

Stark KD, Regula G, Hernandez J, Knopf L, Fuchs K, Morris RS, Davies P, 2006. Concepts for risk-based surveillance in the field of veterinary medicine and veterinary public health: review of current approaches. BMC Health Serv Res 6: 20.

Thrusfield M, 1995. Veterinary Epidemiology 2nd edition. Blackwell Science Ltd.: Oxford. 183-188.

Tumpey TM, Kapczynski DR, Swayne DE, 2004. Comparative susceptibility of chickens and turkeys to avian influenza A H7N2 virus infection and protective efficacy of a commercial avian influenza H7N2 virus vaccine. Avian Dis. 48: 167-176.

van den Berg T, Houdart P, 2008. Avian Influenza outbreak management: action at time of confirmation, depopulation and disposal methods; the 'Belgian experience' during the H7N7 Highly Pathogenic Avian Influenza Epidemic in 2003. Zoonoses Public Health. 55: 54-64.

Vose D, 2000. Risk analysis 2nd edition. John Wiley & Sons Ltd.: New York. 1-418.

Webster RG, Peiris M, Chen H, Guan Y, 2006. H5N1 outbreaks and enzootic influenza. Emerg Infect Dis. 12: 3-8.

World Animal Health Organisation (OIE), Terrestrial Code Manual, Avian Influenza. URL (last update date: 08/07/2005): http://www.oie.int/eng/normes/mmanual/A_00037.htm, (consulted date 27/05/2008).

Appendix

Table 6: Summary for East Flanders of the sensitivities, previous sample sizes, as well as optimal and optimised sample sizes estimation obtained in each risk group in East Flanders

RG	Risk Zone				Non risk zone			
	Se	Initial SS	Optimal SS	Optimised SS	Se	Initial SS	Optimal SS	Optimised SS
Out Chicken	0.52	20	150*	150	0.33	120	260*	250
Out Turkey	0.79	40	40	40	0.13	40	380*	370
Out GD	0	0	0	0	0.13	200	200	200
Out Ratites	0	0	0	0	0	0	0	0
Out OtherP	0	0	0	0	0.07	20	500*	480
In Chicken	0.99	300	80	100	0.74	850	120	120
In Turkey	0	0	0	40	0.09	60	500*	450
In GD	0	0	0	0	0.21	750	1100*	1100
In Ratites	0	0	0	0	0	0	0	0
In OtherP	0.3	200	170*	40	0.99	60	500*	450

*It was impossible to achieve the desired accuracy by sampling every unit in that category, therefore the optimal number of samples to be taken was given; but it could be considered to spread this number of samples over the year to reach the goal.

Table 7: Summary for West Flanders of the sensitivities, previous sample sizes, as well as optimal and optimised sample sizes estimation obtained in each risk group in West Flanders

RG	Risk Zone				Non risk zone			
	Se	Initial SS	Optimal SS	Optimised SS	Se	Initial SS	Optimal SS	Optimised SS
Out Chicken	0.94	40	70*	70	0.77	270	100	100
Out Turkey	0	0	0	0	0.29	60	160*	150
Out GD	0	0	0	0	0	0	34	0
Out Ratites	0	0	0	0	0	0	0	0
Out OtherP	0	0	0	0	0.25	50	190*	180
In Chicken	0.99	360	90	100	0.99	2280	120	120
In Turkey	0.83	60	40	40	0.75	500	50	50
In GD	0	0	0	0	0.15	300	350	340
In Ratites	0	0	0	0	0	0	0	0
In OtherP	0.83	60	50	40	0.1	40	500*	450

*It was impossible to achieve the desired accuracy by sampling every unit in that category, therefore the optimal number of samples to be taken was given; but it could be considered to spread this number of samples over the year to reach the goal.

Table 8: Summary for Antwerpen of the sensitivities, previous sample sizes, as well as optimal and optimised sample sizes estimation obtained in each risk group in Antwerpen

RG	Risk Zone				Non risk zone			
	Se	Initial SS	Optimal SS	Optimised SS	Se	Initial SS	Optimal SS	Optimised SS
Out Chicken	0.78	20	60*	60	0.91	420	90	100
Out Turkey	0	0	0	0	0.11	20	30	30
Out GD	0	0	0	0	0	0	0	0
Out Ratites	0	0	0	0	0	0	0	0
Out OtherP	0	0	0	0	0.06	10	10	10
In Chicken	0.99	220	80*	80	0.98	1450	100	100
In Turkey	0.47	20	30	40	0.1	60	300*	300
In GD	0	0	0	0	0	0	0	0
In Ratites	0	0	0	0	0	0	0	0
In OtherP	0	0	0	0	0.03	10	10	10

*It was impossible to achieve the desired accuracy by sampling every unit in that category, therefore the optimal number of samples to be taken was given; but it could be considered to spread this number of samples over the year to reach the goal.

Table 9: Summary for Limburg of the sensitivities, previous sample sizes, as well as optimal and optimised sample sizes estimation obtained in each risk group in Limburg

RG	Risk Zone				Non risk zone			
	Se	Initial SS	Optimal SS	Optimised SS	Se	Initial SS	Optimal SS	Optimised SS
Out Chicken	0.74	20	100*	100	0.79	290	110	110
Out Turkey	0	0	0	0	0.21	40	40	40
Out GD	0	0	0	0	0.11	100	450	450
Out Ratites	0	0	0	0	0	0	0	0
Out OtherP	0	0	0	0	0.21	40	260	250
In Chicken	0.82	60	60	60	0.7	460	130	130
In Turkey	0	0	0	0	0	0	0	0
In GD	0	0	0	0	0.39	900	100	100
In Ratites	0	0	0	0	0	0	0	0
In OtherP	0	0	0	0	0.28	120	170*	170

*It was impossible to achieve the desired accuracy by sampling every unit in that category, therefore the optimal number of samples to be taken was given; but it could be considered to spread this number of samples over the year to reach the goal.

Table 10: Summary for Wallonia of the sensitivities, previous sample sizes, as well as optimal and optimised sample sizes estimation obtained in each risk group in Wallonia

RG	Risk Zone				Non risk zone			
	Se	Initial SS	Optimal SS	Optimised SS	Se	Initial SS	Optimal SS	Optimised SS
Out Chicken	0.81	20	60*	60	0.84	200	90	100
Out Turkey	0	0	0	0	0.12	20	30	30
Out GD	0	0	0	0	0.47	500	80	400
Out Ratites	0.97	40	40	40	0.91	0	80	80
Out OtherP	0	0	0	0	0.23	40	220*	200
In Chicken	0.48	20	150*	150	0.83	600	110	110
In Turkey	0	0	0	0	0.27	100	170*	200
In GD	0	0	0	0	0.36	700	600	600
In Ratites	0	0	0	0	0.45	200	170	160
In OtherP	0	0	0	0	0.2	70	240*	240

*It was impossible to achieve the desired accuracy by sampling every unit in that category, therefore the optimal number of samples to be taken was given; but it could be considered to spread this number of samples over the year to reach the goal.

CHAPTER IV: EVALUATION OF THE BELGIAN SURVEILLANCE PROGRAMME FOR BOVINE BRUCELLOSIS THROUGH SCENARIO TREE SIMULATIONS

Adapted from Epidemiol. et Santé Animale, 55, 1-6, Epub 2009.

Evaluation du programme de surveillance pour la Brucellose et la Leucose Bovine enzootique en Belgique.

Welby S, Letellier C, Fretin D, Hooyberghs J, Vanholme L, Godefroid J, Van der Stede Y.

Abstract

Having maintained the bovine brucellosis officially free status for at least 5 consecutive years, an evaluation of the Belgian brucellosis surveillance program in compliance with the OIE and EU legislation (Council Directive 64/432/EEC) was performed. Scenario trees were used to analyse the bovine brucellosis surveillance data from 2007-2008, in order to evaluate the on-going surveillance components within this program and determine how this program could be optimised. Differential risks of infection or detection partitioned the population structure in infection and detection category risk nodes. The current surveillance system and the impact of reducing surveillance in certain risk groups on the whole surveillance system sensitivity were simulated together stochastically. Results demonstrated that a significant decrease of the total number of samples was feasible and yet maintain a 99% confidence level of disease detection. The total sensitivity for an early detection program would be importantly increased if surveillance activities were targeted towards passive surveillance meaning reinforcing the legal obligation of declaration, reporting and testing of abortions by the farmers and veterinarians. In addition the continuation of testing a certain number of traded animals (within BE) as well as imported cattle remains of utmost importance.

INTRODUCTION

Bovine brucellosis (BB), caused by Brucella abortus (a gram negative bacteria), is a notifiable disease according to the World Organisation of Animal Health (OIE) list (OIE, 2011a). The burden of this highly infectious disease is measured by its economical and public health impact. BB is also important disease for livestock and the most frequent isolated species are Brucella abortus in cattle, Brucella suis in pigs and Brucella melitensis in small ruminants.

The main clinical symptoms in cattle are abortions and ‘abortion storms’ may occur in a naïve population. Humans might get infected by the consumption of contaminated raw milk and/or meat, or by close contact with infected cattle and their secretions. In humans flue like symptoms can be observed and may remain for several months. In some circumstances, infection can lead to chronic liver and spleen lesions as well as infertility in men. Cattle mainly get infected by secretions of contaminated animals, mainly after calving. Brucella bacteria may survive for months in the environment if

the conditions are favourable, such as high humidity and cold weather. Consequently these bacteria can remain a constant source of infection on holdings and calves can get infected following suckling or drinking of raw milk from a contaminated animal (OIE, 2011b; Wilkinson,1993).

Within the European Union (EU), at the time of the study, BB due to Brucella abortus was eradicated in most of the 27 Member States (MS) (AT, BE, CZ, DE, DK, FI, FR, IE, LU, NL, SE, SI, SK, PL) while in some other MS only specific areas were declared free of brucellosis (IT, PT, ES, and UK). This means that within the EU the disease is still prevalent in 10 MS (2003/467/EC).

The OIE and the European Commission requires a compulsory surveillance program in order to either eradicate or obtain/maintain the officially 'freedom of disease' status in each MS, and allow trade (64/432/EC; OIE, 2011c). In order to monitor a potential re-emergence of the disease, and allow trade, a continuous serological testing scheme is carried out in BE according to EU Directive 64/432/EEC (Godefroid et al., 2002; Godefroid and Kasbohrer, 2002; OIE, 2011b, c; Seleem, 2010).

In Belgium, the surveillance program for BB consists in active and passive surveillance components that run in parallel (MB, 1978). The active surveillance component consists of antibody detection in serum and milk, in beef and dairy cattle respectively. This serological (serum) screening is carried on all imported cattle from officially free, or non-officially free countries, traded cattle in Belgium and on 33% of all Belgian beef cattle farms during the winter screening (November-March) ensuring that the whole population is tested every 3 years. All imported female cattle, above 24 month of age, originating from non-officially free countries, whether introduced in dairy or beef herds, are serologically tested for three consecutive years during the winter screening. The milk testing is carried out on dairy herds at least 4 times a year with the Milk ring test (MRT) test. The passive surveillance consists of clinical surveillance and reporting all abortions cases by the field veterinarians. These abortions are further investigated in the regional laboratories via coloration and primo culture (OIE, 2011c).

Using these different surveillance components, Belgium has obtained the officially free status for brucellosis since 2003 (2003/467/EC). Council Directive 64/432/EC determines that a Member State with an officially free status (for BB) for 5 consecutive years is allowed to re-evaluate its surveillance programme in relation to

frequency of sampling and sample size (64/432/EC). Following a demand of the Belgian Federal Agency for the Safety of the Food Chain (FASFC), the current study aimed at evaluating the current Belgian surveillance programmes of BB, in order to estimate the relative sensitivity of each surveillance component and identify the most optimal surveillance system protocol.

For this study, quantifying the sensitivity of the surveillance programme to substantiate freedom of disease, assuming 100% specificity, was the major aim.

The use of scenario trees has been chosen to conduct this study and accounts for number of tested samples, diagnostic tests applied, the minimum design prevalence to be detected and the differential risk of infection or detection. This method offers interesting tools to quantify the surveillance system sensitivity and has proven its efficacy in the past (Frossling et al., 2009; Hadorn et al., 2008; Knight-Jones et al., 2010; Martin et al., 2007a,b; Stark et al., 2006; Welby et al., 2010).

MATERIALS AND METHODS

Scenario tree for surveillance of bovine brucellosis in Belgium

A scenario tree was created for BB as described in detail by Martin et al. (Martin et al., 2007a). All factors interfering with the probability of infection or detection were represented in the same tree by category, infection or detection nodes respectively which further would define the limbs of the tree (risk groups) for each category node combination. Category nodes (either of infection or detection), infection nodes and detection nodes, as described here below, have their respective parameters in order to obtain the probabilities of detection at animal level (ASe) and herd level (HSe) in each limb of the tree, given the country would be infected at a certain design prevalence. The combination of the obtained individual HSe enables the computation of the total surveillance sensitivity (SSe) (Figure 1). For illustration purpose only one possible risk node combination is illustrated in this figure though all possible combinations were accounted for in the analysis.

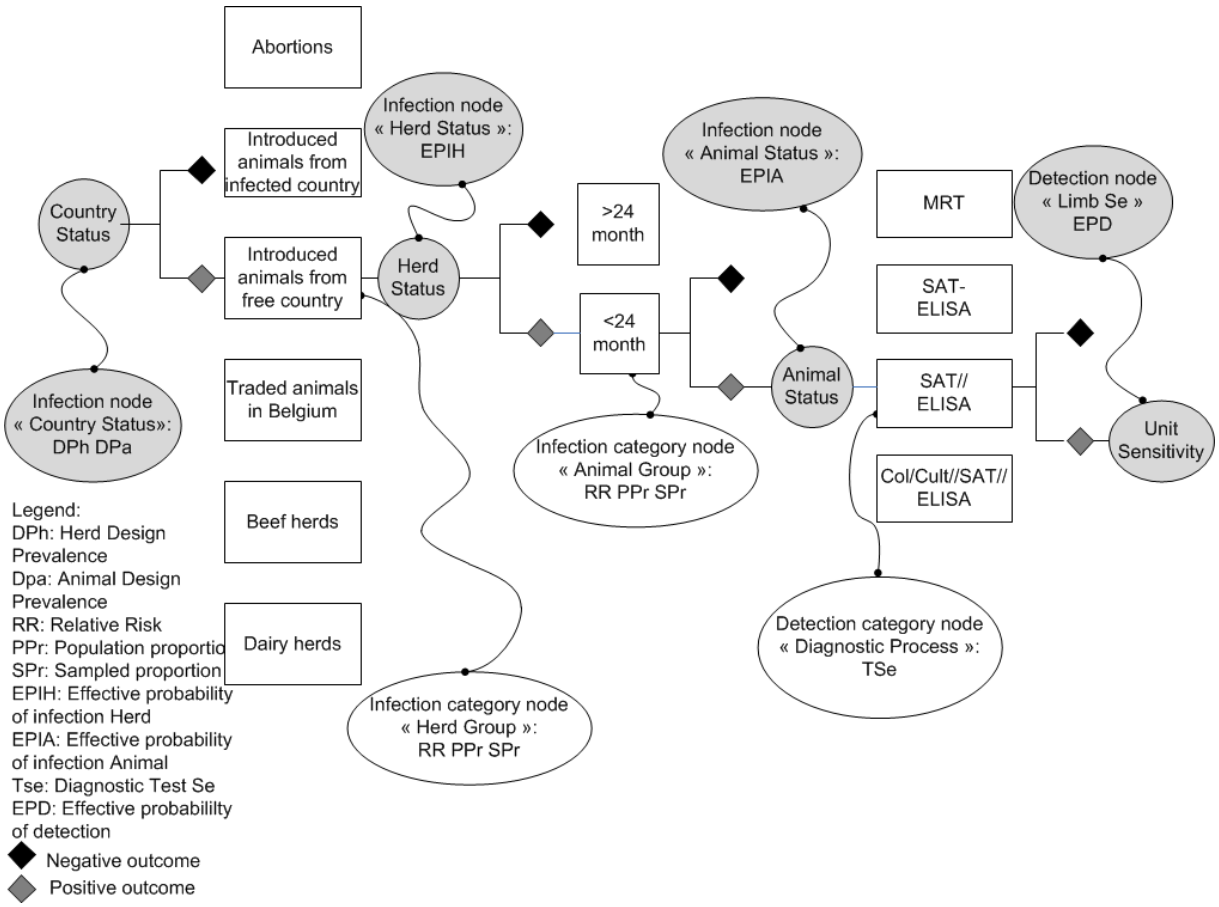


Figure 1: Scenario tree for Belgian bovine brucellosis surveillance (only one combination is illustrated for clarity purpose).

Data sources for each category, infection and detection node

Expert elicitation was used to estimate the relative risk that could partition the population towards the differential risk of infection or detection. Different Belgian stakeholders as well as experts from the National Reference Laboratory (CODA-CERVA), veterinary syndicates (UPV), farm syndicates (Boerenbond, FWA), FASFC and regional veterinary laboratories DGZ and ARSIA were consulted for this purpose. This enabled to identify the category nodes of the tree.

All risk groups were assumed to be independent. The first node defined is an infection node “Country Status” with its targeted design prevalence at herd level (DPh) and animal level (DPa). Following the infection node country status, the infection category node “Herd Group” was represented. This node represents the different subpopulations in the Belgian cattle population namely: i) Abortions (Ab), ii)

Imported cattle from a brucellosis non-free country (IR), iii) Imported cattle from a brucellosis free country (IF), iv) Traded cattle in Belgium (Tr), v) Beef herds (WS), vi) Dairy herds (Da).

The category node 'Herd Group', followed by its respective "Herd Status" infection node, was further subdivided into an infection category node "Animal Group" for which 2 different age categories (<24 month of age and ≥ 24 month of age) were considered according to the differential risk of infection in each category.

Subsequently, a next infection node at animal level namely the "Animal Status" could be defined. The diagnostic process for BB is represented in the detection category node. The parameter characterizing this node is represented by the diagnostic process sensitivity (TSe), which accounted for each diagnostic test used in the process namely: i) For dairy cattle, the milk ring test (MRT), ii) For traded cattle in Belgium and cattle tested during the WS, the Slow Agglutination Test with EDTA (SAT-EDTA) (OIE, 2011c) and commercial indirect ELISA test (Synbiotics®) were performed serially if the SAT-EDTA test was positive. If this indirect ELISA test was positive, the sample(s) were tested with an in house indirect ELISA carried out by the National Reference Laboratory (NRL) in Belgium (SAT-ELISA) (EFSA, 2006), iii) For imported cattle SAT-EDTA test and commercial indirect ELISA were performed in parallel. If at least one of these assays is positive the in house indirect ELISA is carried out by the NRL (SAT//ELISA), iv) For abortive tissues, coloration (Ziehl-Nielsen) and a primo culture (primary culture for Brucella spp.) are performed in parallel. In addition the SAT-EDTA test and a commercial indirect ELISA is performed on the serum of the aborted animal (COL/CULT//SAT//ELISA).

The DPh is obtained from the legal requirements defining the legal minimum detection level of disease according the EU and OIE. For BB, the DPh is set at 0.2%, maximum herd prevalence (64/432/EC; OIE, 2011c; MB, 1978). For the DPa, no minimum legal obligations were prescribed but a 30% DPa in a naïve population (after eradication) was assumed. This estimate was based on expert opinion and observed minimum within-herd prevalence for other epizootic disease with similar epidemiology as BB. In order to measure the impact of a fixed DPa, a sensitivity analysis was performed for this parameter.

Regarding the category nodes, corresponding parameters for each representative subpopulation (PPr), sampled population (SPr) as well as relative risk (RR) were

attributed. Population proportions and sampled data were obtained from the 2007-2008 Belgian Animal Identification and Registration system (SANITEL) of the FASFC as well as from the two official and accredited regional veterinary laboratories in Belgium (ARSIA, DGZ) involved in official animal disease surveillance. In order to estimate the real proportion (PrP) of abortive females in the Belgian cattle population, an abortion rate of 3 to 5 % was assumed, corresponding to an average minimum of 1 abortion per Belgian cattle herd per year. In 2008 around 4,056 abortions were registered.

For the diagnostic tests sensitivities, literature review (EFSA, 2006; Gall and Nielsen, 2004; Godefroid et al., 2010) was carried out and expert opinion was elicited as well. Appropriate distributions were fitted to the data that account for the uncertainty and variability in the data.

Parameters obtained and computed for each category, infection and detection nodes

The different parameters obtained for each category node described above, namely the PPr and RR enabled the calculation of an adjusted risk of infection (AR) for each "Herd Group" by (Eq. 1);

$$\mathrm{AR}_{\mathrm{HerdGroup}i} = \frac{\mathrm{RR}_{\mathrm{HerdGroup}i}}{\sum \mathrm{RR}_{\mathrm{HerdGroup}i} * \mathrm{PPr}_{\mathrm{HerdGroup}i}} \text{ (Eq.1)}$$

The infection status at herd level (="Herd Status" infection node) is characterised by the EPIH and takes into account the adjusted risk of infection as well as the design prevalence at herd level as shown in (Eq. 2);

$$\mathrm{EPIH}_{\mathrm{HerdGroup}i} = \mathrm{DPh} * \mathrm{AR}_{\mathrm{HerdGroup}i} \text{ (Eq. 2)}$$

The same was calculated at animal level where for each "Animal Group" and for each age category (<24 month and ≥ 24 months) with the corresponding differential risk of infection for the proportion of cattle in each category (Eq. 3);

$$\mathrm{AR}_{\mathrm{AnimalGroup}i} = \frac{\mathrm{RR}_{\mathrm{AnimalGroup}i}}{\sum \mathrm{RR}_{AnimalGroupi} * \mathrm{PPr}_{\mathrm{AnimallGroup}i}} \text{ (Eq. 3)}$$

The infection status at animal level (="Animal Status" infection node) is defined by its respective probability of infection (EPIA) and derived from the adjusted risk and design prevalence at animal level as shown in (Eq. 4);

$$\mathrm{EPIA}_{\mathrm{AnimallGroup}i} = \mathrm{DPa} * \mathrm{AR}_{\mathrm{AnimallGroup}i} \text{ (Eq. 4)}$$

Using the diagnostic test sensitivity (TSe) used and the fraction of population sampled (SPr) with that specific test, the effective probability of detection (EPD) was computed for each unit sampled in each limb of the tree (Eq. 5).

$$\mathrm{EPD}_{\mathrm{AnimalGroup}i} = \mathrm{SPr}_{\mathrm{HerdGroup}i} * \mathrm{SPr}_{\mathrm{AnimallGroup}i} * \mathrm{TSe}i \text{ (Eq. 5)}$$

For each limb of the tree (defined by each possible risk group combination), respective HSe and ASe were then estimated. These calculations were based on the parameters obtained above.

For each ASe of each risk group identified, the binomial sampling approach was considered if the fraction of population sampled was <10% (Ab, (Eq. 6)) while the hypergeometric sampling approach was used in the estimation process if the sampled population was a fraction (>10%) of the total population (WS and Da) (Eq. 7) and the exact sampling method was used if the whole population was sampled (IR, IF and Tr, shown in Eq. 8), as described by Martin et al. (Martin, et al., 2007a) NAnimalGroupi characterized the total number of animals in that branch and nAnimalGroupi is the number of animals sampled.

$$\mathrm{ASe}_{\mathrm{AnimalGroup}i} = 1 - (1 - \mathrm{EPD}_{\mathrm{AnimalGroup}i} * \mathrm{EPIA}_{\mathrm{AnimalGroup}i})^{\mathrm{n}_{\mathrm{AnimallGroup}i}} \quad \text{(Eq. 6)}$$

$$\mathrm{ASe}_{\mathrm{AnimalGroup}i} = 1 - (1 - \mathrm{EPD}_{\mathrm{AnimalGroup}i} * \left(\frac{\mathrm{n}_{\mathrm{AnimallGroup}i}}{\mathrm{N}_{\mathrm{AnimallGroup}i}}\right)^{\mathrm{EPIA}_{\mathrm{AnimalGroup}i} * \mathrm{N}_{\mathrm{AnimallGroup}i}} \text{ (Eq. 7)}$$

$$\mathrm{ASe}_{\mathrm{AnimalGroup}i} = 1 - (1 - \mathrm{EPD}_{\mathrm{AnimalGroup}i})^{\mathrm{EPIA}_{\mathrm{AnimalGroup}i} * \mathrm{N}_{\mathrm{AnimallGroup}i}} \text{ (Eq. 8)}$$

The same was calculated for each HSe: for abortions a binomial distribution was used (Eq.9) as the fraction of herds reporting abortions was only 7%, while for beef herds a hypergeometric approach was used (Eq. 10). For imported cattle, national traded cattle, as well as dairy herds the exact method was used (Eq. 11) because all herds were tested. NHerdGoupi characterized the total number of herds in that branch and nHerdGroupi is the number of herds sampled.

$$\mathrm{HSe}_{\mathrm{HerdGroup}i} = 1 - (1 - \mathrm{ASe}_{\mathrm{AnimalGroup}i} * \mathrm{EPIH}_{\mathrm{HerdGroup}i})^{\mathrm{n}_{\mathrm{HerdGroup}i}} \text{ (Eq. 9)}$$

$$\mathrm{HSe}_{\mathrm{HerdGroup}i} = 1 - (1 - \mathrm{ASe}_{\mathrm{AnimalGroup}i} * (\frac{\mathrm{n}_{\mathrm{HerdGroup}i}}{\mathrm{N}_{\mathrm{HerdGroup}i}}))^{\mathrm{N}_{\mathrm{HerdGroup}i} * \mathrm{EPIH}_{\mathrm{HerdGroup}i}} \text{ (Eq. 10)}$$

$$\mathrm{HSe}_{\mathrm{HerdGroup}i} = 1 - (1 - \mathrm{ASe}_{\mathrm{AnimalGroup}i})^{\mathrm{N}_{\mathrm{HerdGroup}i} * \mathrm{EPIH}_{\mathrm{HerdGroup}i}} \text{ (Eq. 11)}$$

Finally, the whole SSe was estimated based on the following formula (Eq. 12);

$$\mathrm{SSe} = 1 - \prod(1 - \mathrm{HSe}_{\mathrm{HerdGroup}i}) \text{ (Eq. 12)}$$

Scenarios simulation

The scenario tree was designed in an Excel spread sheet (Excel, 2007). Firstly, the outputs for the HSe and ASe were estimated through hypergeometric simulations, with 10,000 iterations per simulation, using ModelRisk 3.0® (Vose, 2010). Taking the uncertainty and variability of parameters into account by fitting appropriate parameter distributions, allowed to account for the range of possible situations.

First the current surveillance system was evaluated.

Secondly, the impact on the total surveillance system sensitivity if one or more surveillance components were reduced or skipped was evaluated.

Thirdly, the monthly posterior probability of freedom (PostFreemonth) was calculated and based on the surveillance system sensitivity observed (simulated) that specific month, as well as the prior probability of introduction (PIntro) and probability of infection (PriorInfmonth) (Eq. 13, 14). The prior probability of infection was considered at 0.5 for the first month and the probability of introduction was held constant (0.5).

$$\mathrm{PostFree}_{\mathrm{month}} = \frac{1-\mathrm{PriorInf}_{\mathrm{month}}}{1-\mathrm{PriorInf}_{\mathrm{month}} * \mathrm{HSe}i} \text{ (Eq. 13)}$$

$$\mathrm{PriorInf}_{\mathrm{month}} = \mathrm{PIntro} + \mathrm{PostInf}_{\mathrm{month}-1} - \mathrm{PIntro} * \mathrm{PostInf}_{\mathrm{month}-1} \text{ (Eq. 14)}$$

Sensitivity analysis

A sensitivity analysis was carried out, (using @risk 4.5®, Palisade) to validate the use of the given parameters and to evaluate the most influential input parameter on the SSe.

In order to determine the most relevant RR for each risk category node, different simulations were carried out. A first analysis, data not shown, was conducted to estimate the impact of choosing RR according to expert opinion, RR equal to 1 in all risk category nodes, very high RR and unrealistic RR for each category node. Finally the most appropriate RR of a given node, taking into account the respective population proportion in that given node separately and the risk perception of the stakeholders were estimated using the SOLVER option of Excel (2007).

In a second step, separate univariate analysis for each HSe were conducted to estimate the influence of each input parameter on the HSe, looking at the rank order correlation statistic ρ. The higher the ρ, the strongest the correlation was between the selected input variable and the output.

Finally in a third step to assess the impact of the different HSe on the CSe, the rank order correlation statistic ρ was used as well to estimate the most influential HSe when changes in sample sizes were simulated. Results

Data sources for each category, infection and detection node

Table 1 illustrates each risk node with the respective parameters PPr, RR. The population proportions were derived from the Belgian Animal Identification and Registration system (SANITEL) of the FASFC. The RR were obtained following multiple iterations trials which took in account expert opinion, very high RR, completely wrong RR and RR equal to 1. Finally the RR which reflected the better the Belgian situation was estimated using Solver in Excel. The same RR was given to IR and Ab, as both constituted a very high risk of disease introduction.

Table 1: Population proportions (PrP) and relative risks (RR) distributions (Distr) for each category risk node identified in the scenario tree for Bovine Brucellosis in Belgium.

Herd level	$N_{HerdGroup}$	PPr Distr	RR Distr
Ab	36000	Normal(0.369;0.037)	Pert(270;300;330)
IR	114	Normal(0.001;0.000)	Pert(270;300;330)
IF	2415	Normal(0.025;0.002)	Pert(90;100;110)
Tr	23000	Normal(0.236;0.024)	Pert(45;50;55)
WS	26000	Normal(0.267;0.027)	Uniform(0;1;1)
Da	10000	Normal(0.103;0.010)	Uniform(0;1;1)
Animal level	$N_{AnimalGroup}$	**PPr Distr**	**RR Distr**
Ab<24month	0	Normal (0.000;0.000)	Uniform(0;1;1)
Ab>24 month	4056	Normal (1.000;0.1000)	Pert(2.5;3;3.5)
IR<24 month	6471	Normal (0.905;0.090)	Uniform(0;1;1)
IR>24 month	683	Normal (0.095;0.009)	Pert(2.5;3;3.5)
IF<24 month	39748	Normal (0.752;0.075)	Uniform(0;1;1)
IF>24 month	13141	Normal (0.248;0.025)	Pert(2.5;3;3.5)
Tr<24 month	402010	Normal (0.677;0.068)	Uniform(0;1;1)
Tr>24 month	191974	Normal (0.323;0.032)	Pert(2.5;3;3.5)
WS<24 month	1075800	Normal (0.489;0.049)	Uniform(0;1;1)
WS>24 month	1124200	Normal (0.511;0.051)	Pert(2.5;3;3.5)
Da<24 month	244500	Normal (0.489;0.049)	Uniform(0;1;1)
Da>24 month	255500	Normal (0.511;0.051)	Pert(2.5;3;3.5)

Representative population data ($N_{HerdGroup}$, $N_{AnimalGroup}$); representative population distribution (PPr Distr(Normal(mean, sd)) and relative risk distribution (RR Distr(Pert(minimum, most likely, maximum); Uniform(minimum, most likely, maximum)); Traded cattle within Belgium (Tr); Dairy cattle (Da); Winter screening surveillance (WS); Import from non-free BB countries (risk countries) (IR); import from free BB countries (IF); Abortions (Ab), above (>24mth) and below 24 month (<24mth).

Parameters obtained and computed for each category, infection and detection nodes

Table 2 illustrates the SPr at animal and herd group level, corresponding to the diagnostic TSe used for each respective risk node as the ASe and HSe resulting from each simulation.

Table 2: Sampled population proportion (SPr), Test sensitivity (TSe), calculated estimated Animal sensitivity (ASe) and Herd sensitivity (HSe) estimations for each risk group.

Risk Group	nAnimal	nHerd	TSe	Ase	HSe
Ab<24mth	0	2424	Col/Cult (Pert(0.95;0.98;0.99))	0.009 (0.008-0.012)	0.11 (0.073-0.160)
Ab>24mth	4056		Col/Cult (Pert(0.95;0.98;0.99))	0.029 (0.03-0.03)	0.286 (0.218-0.393)
IR<24mth	6471	114	SAT//ELISA(Pert(0.9;0.92;0.95))	1(1-1)	1(1-1)
IR>24mth	683		SAT//ELISA(Pert(0.9;0.92;0.95))	1(1-1)	1(1-1)
IF<24mth	39748	2415	SAT//ELISA(Pert(0.9;0.92;0.95))	1(1-1)	1(1-1)
IF>24mth	13141		SAT//ELISA(Pert(0.9;0.92;0.95))	1(1-1)	1(1-1)
Tr<24mth	402010	23000	SAT-ELISA(Pert(0.8;0.86;0.9))	1(1-1)	1(1-1)
Tr>24mth	191974		SAT-ELISA(Pert(0.8;0.86;0.9))	1(1-1)	1(1-1)
WS<24mth	0.000	8600	SAT-ELISA(Pert(0.8;0.86;0.9))	0(0-0)	0(0-0)
WS>24mth	1124200		SAT-ELISA(Pert(0.8;0.86;0.9))	1 (0.99-1)	0.154 (0.115-0.222)
Da<24mth	244500	10000	MRT(Pert(0.85;0.89;0.95))	0.768 (0.651-0.887)	0.210 (0.129-0.332)
Da>24mth	255500		MRT(Pert(0.85;0.89;0.95))	1 (0.999-1)	0.987 (0.940-1)

Sampled population data ($n_{AnimalGroup}$, $n_{HerdGroup}$); test sensitivities (TSe) (Pert(minimum, most likely, maximum); risk group sensitivity obtained at animal (ASe) and herd (HSe) level (minimum and maximum values are between brackets); Traded cattle within Belgium (Tr); Dairy cattle (Da); Winter screening surveillance (WS); Import from non-free BB countries (risk countries) (IR); import from free BB countries (IF); Abortions (Ab), above (>24mth) and below 24 month (<24mth).

The higher sensitivities were obtained for the imported and traded (within BE) cattle. The mean, minimum and maximum values across the 10,000 iterations history process are illustrated for the ASe and HSe in each risk group. Despite the poor sensitivity of the MRT, systematic testing of all dairy herds resulted in a high HSe. As the number of cattle tested within herd is high, WS had high ASe, but due to the fact that only 1/3 of the total cattle herd population is tested in that group, low HSe was observed for that herd group. Due to the limited number of abortions tested this risk group had very poor ASe and HSe, despite the very high sensitivity of diagnostic test used.

Scenarios simulation

The whole surveillance system sensitivity was equal to one. Results showed the impact of stopping surveillance in one of the risk groups, namely; i) Only testing Ab, IR, IF, Tr, Da, leaving out WS ii) Only testing Ab, IR, IF, Tr, WS, leaving out Da, iii)

Only testing Ab, IR, IF, Tr, leaving out Da and WS, iv) Only testing Ab, Tr, WS, Da, leaving out IR, IF, v) Only testing Ab leaving out IR, IF, Tr, WS, Da.

None of these modifications turned out to be influential on the whole surveillance sensitivity which remained one. This indicated that reduction in the total number of samples was allowed and yet maintain the same sensitivity of detection. However to obtain an optimal detection level, a targeted approach is required.

Figure 2 illustrates the monthly evolution of the posterior probability of freedom in each risk group. A sufficient guarantee with relation to the posterior probability of freedom is obtained if imported cattle from free or non-free countries are tested (Figure 2).

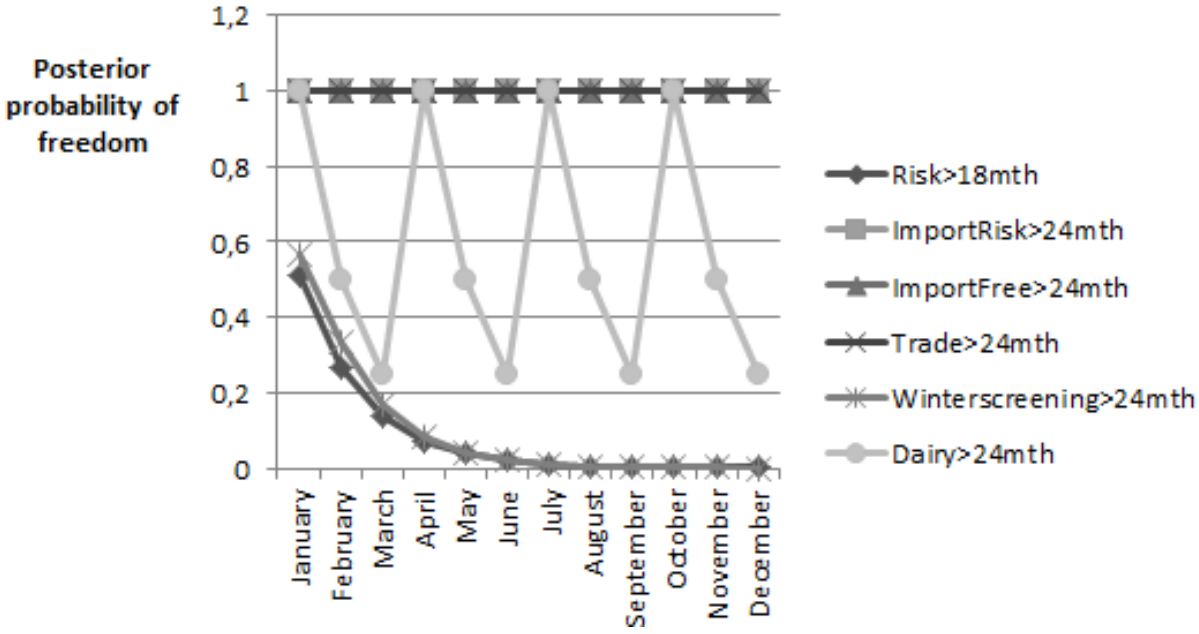


Figure 2: Monthly evolution of Posterior Freedom Probability for bovine brucellosis in Belgium.

Testing the WS group and Da group is of limited value as it only happens once in a year (WS), or 4 times a year (Da). The current number of abortions tested per year is of very poor value to insure a posterior probability of freedom.

Sensitivity analysis

The sensitivity analysis showed that the Ab surveillance component is the most influential parameter on the SSe, especially when sample size changes were simulated. A 10% increase in the sensitivity was observed when the number of reported was doubled!

The input parameter affecting most the individual HSe was the diagnostic test used namely the COL/CULT//SAT//ELISA and the MRT (ρ=0.73-0.68). The RR attributed to Ab, and Tr, had a small influence (ρ=0.35-0.28). All the other input parameters had only a very poor (ρ < 0.01) influence. When a pert distribution reflecting a large variation around the DPa (min =0.05, most likely=0.3 and max=0.7), this parameter had only little influence on the HSe and SSe (ρ<0.01), compared to all the other input parameters in the model.

DISCUSSION

The results in this study showed that the overall surveillance system sensitivity for BB, mainly based on serological screening, in Belgium was very high. This is probably due to the large amount of herds processed within the different risk groups as shown in Table 1. Moreover, the impact of suppressing surveillance in a specific risk group did not have huge impact on the whole surveillance system sensitivity. Testing at import, or following introduction of cattle in a herd seem to play a crucial role with regard to proving freedom of disease and for early detection, as shown in Table 2 and Figure 2. Indeed, in Belgium around 450,000 to 500,000 trade 'activities' (purchases) are recorded each year and allows a systematic control on a daily basis. Interestingly, despite the poor TSe of the MRT, testing of dairy cattle turns out to be quite effective (Table 2), partly due to the large amount of herds tested and the low RR attribute to this risk group, but remains of limited value as it is only carried out 4 times a year (Figure 2). For WS, the ASe is quite high, this is because during WS in the tested herds, nearly all animals are sampled (one stage sampling). Of course because only a fraction (33%) of the herds is tested once per year, the HSe is low as well as the posterior probability of freedom (Table 2, and Figure 2). Strikingly, through this study it was obvious that the ASe and HSe in the Ab component were very low. The sensitivity analysis showed however that increasing the number of samples taken in that risk group would have a high impact on the HSe in that group. The 4,056 abortions registered in 2008 suggest that underreporting for this 'sampled' population fraction is substantial and that reporting of abortions by field veterinarians should be stimulated.

It should be noted that the output of the scenario tree simulations depends on the input parameters which have been set up in the model. This is why in the present study a thorough sensitivity analysis was conducted on each input parameter. From that sensitivity analysis it became clear that the expected within herd design prevalence had no influence on the output. This means that results of this model are quite robust despite the uncertainty around this parameter (0.05-0.7). Interestingly the TSe of testing abortions was seen as the most influential parameter, and further underlines the importance of increasing the efforts to rise up these values either by improving the diagnostic methods or the number of samples. The value of the MRT was also influential on the results; this can be easily explained by its large range. Thus an alternative testing protocol for dairy herds is advisable. Finally the RR attributed to abortions and purchased animals was also influential. It must be stated however that in this study the most optimal value was estimated following the numerous sensitivity analysis, thus it can be assumed that the chosen values truly reflected the 'real life' situation (Table 1). However obtaining more empirical data for these values would be interesting in the future.

Nevertheless these models provide insight and understanding on how to improve the efficiency of a surveillance program.

Through this study new questions have arisen. Although it was found that focusing the surveillance on reporting and investigation of abortions would constitute an increased confidence level towards probability of freedom, it will be hard to put into practice due to farmer's difficulty to notify an abortion. Indeed, some distrust to the competent authority and the fear of economic consequences as well as the ethical impact of a notification should be considered (Elbers et al., 2010). Therefore it would be useful to evaluate how this notification could be improved, e.g. through information campaigns, or by stressing the legal obligation by the veterinarians, or by providing a financial compensation to those farmers reporting their abortions. This certainly will improve the passive surveillance not only for BB but also other infectious diseases possibly leading to abortion such as e.g. listeria, leptospirosis, bovine viral diarrhea and Q-fever. Recently an 'abortion' protocol for cattle in Belgium is installed in the 2 official regional veterinary laboratories and allows the simultaneous detection of various pathogens such as Brucellosis, Listeriosis, BVD, BTV, Q Fever, Neospora, IBR, Leptospirosis, Listeriosis, Mycosis (champignons),

other pathogens (E coli, pseudomonas,...). Furthermore a random selection of 15,000 cattle from all herds which did not declare any abortions during the past year were investigated for BB during WS by the means of SAT ELISA test.

With regard to the diagnosis of imported cattle it is of utmost importance to implement the most optimal testing protocol, the time to seroconvert after infection in bovines with Brucella abortus, varies between 3-4 weeks (in case of an acute infection) to sometimes 2 years (in case of a latent infection). In order to optimize the surveillance programme, a protocol taking into account this physio-pathological lag period must be applied. Another important issue to take into consideration is the time lag before the results of tested traded cattle are available. Indeed, actually almost no quarantine or biosecurity measures are respected at purchase. The cattle are sampled upon arrival at the farm of destination, and cattle remain there until receipt of the laboratory results. During that lag phase the infection could probably further spread within a herd or between herds if no quarantine or biosecurity measures are respected. Therefore, the estimation of the RR should also be based on this knowledge. That's why high (and even higher) estimations of RR for introduced cattle are to be considered. The opinion of different stakeholders, related to the import and export movements of cattle was useful in order to have more accurate quantification of the RR for this specific risk category of bovines.

CONCLUSION

Following this study, stakeholders agreed on the importance of testing imported cattle as well as abortions. To ensure a minimal awareness amongst farmers and labs as well as to maintain minimal technical resources and technician expertise in case of a re-emergence of the disease, it was agreed that testing of a minimum number of samples during WS and Tr should be continued. The surveillance program in Belgium was modified in 2010 (FASFC, 2011). Within the modified surveillance program the number of samples was estimated to be around 30,000. The reporting of abortion was strongly stimulated, via the abortion protocol. Via this protocol the aim was to reach an annual number of reports of 8,000 abortions instead of the 4,000. Furthermore a random selection of 15,000 cattle from all herds which did not declare any abortions during the past year were investigated for BB during

WS by the means of SAT ELISA test. In addition, it was recommended to test 8,500 cattle during purchase (randomly selected in regional laboratories) and 8,500 cattle from intra community trade from officially, as well as all cattle from intra community trade from non-officially free countries.
The new surveillance program led to the detection one case in 2010 (OIE, 2011d) following abortion investigation. This finding corroborated the usefulness of the present study.

To conclude, this study evaluated, as best as possible, the real life situation of the current brucellosis surveillance program. The results of the analysis provide insight on how the surveillance can more efficiently be redesigned. Also this study constitutes a valuable tool to provide a required confidence level to the surveillance system in order to respect the international standard requirements.

REFERENCES

Animal Health World Organisation (OIE), 2011. Chapter 11.3, Bovine Brucellosis. In Terrestrial Animal Health Code. URL (Last update date: 2011): http://web.oie.int/eng/normes/mcode/en_chapitre_1.11.3.htm (Consulted date: 04/05/2011).

Animal Health World Organisation (OIE), 2011. OIE Listed diseases. URL(Last update date: 2011):http://www.oie.int/en/animal-health-in-the-world/oie-listed-diseases-2011/ (Consulted date: 06/05/2011).

Animal Health World Organisation (OIE), 2011. Animal disease: Technical specifications: Bovine Brucellosis. URL(Last update date: 2011): http://www.oie.int/fileadmin/Home/fr/Media_Center/docs/pdf/Disease_cards/BCLS-FR.pdf (Consulted date: 04/05/2011).

Animal Health World Organisation (OIE), 2011. WAHID Interface, Animal health information. URL(Last update date: 2011): http://web.oie.int/wahis/public.php?page=weekly_report_index&admin=0 (Consulted date: 06/05/2011).

European Commission (EC), 2003. Commission Decision 2003/467/EC establishing the official tuberculosis, brucellosis, and enzootic-bovine-leukosis-free status of certain Member States and regions of Member States as regards bovine herds (last amended by Commission Decision 2010/695/EU). Official Journal of the European Union L156, 74-78.

European Commission (EC), 1964. Council Directive 64/432/EEC on animal health problems affecting intra-Community trade in bovine animals and swine. Official Journal of the European Union L121, 1977-2012.

European Food Safety Agency (EFSA), 2006. Performance of Brucellosis Diagnostic Methods for Bovines, Sheep, and Goats. EFSA Journal 432, 1-44.

Elbers, A.R.W., Gorgievski-Duijvesteijn, M.J., Zarafshani, K., Koch, G., 2010. To report or not to report: a psychosocial investigation aimed at improving early detection of avian influenza outbreaks. Rev. Sci. Tech. Off. Int. Epiz. 29, 435-449.

Federal Agency for the Security of the Food Chain (FASFC), 2011. Scientific Committee: Evaluation de l'allègement de la surveillance de la brucellose et de la leucose bovines et propositions pour un nouveau programme de surveillance d'autres maladies bovines (dossier Sci Com 2009/25). [on line] (2011) http://www.favv-afsca.fgov.be/comitescientifique/avis/_documents/Addendumavisrapide26-2009_DOSSIER2009_25.pdf [consulted 6 May 2011].

Frossling, J., Agren, E.C., Eliasson-Selling, L., Lewerin, S.S., 2009. Probability of freedom from disease after the first detection and eradication of PRRS in Sweden: scenario-tree modelling of the surveillance system. Prev. Vet. Med. 91, 137-145.

Gall, D., Nielsen, K., 2004. Serological diagnosis of bovine brucellosis: a review of test performance and cost comparison. Rev. Sci. Tech. Off. Int. Epiz. 23, 989-1002.

Godfroid, J., Nielsen, K., Saegerman, C., 2010. Diagnosis of brucellosis in livestock and wildlife. Croat. Med. J. 51, 296-305.

Godfroid, J., Kasbohrer, A., 2002. Brucellosis in the European Union and Norway at the turn of the twenty-first century. Vet. Microbiol. 90, 135-145.

Godfroid, J., Saegerman, C., Wellemans, V., Walravens, K., Letesson, J.J., Tibor, A., Mc Millan, A., Spencer, S., Sanna, M., Bakker, D., Pouillot, R., Garin-Bastuji, B., 2001. How to substantiate eradication of bovine brucellosis when aspecific serological reactions occur in the course of brucellosis testing. Vet. Microbiol. 90, 461-477.

Moniteur Belge (MB), 1964. Arrêté royal du 06.12.1978 relatif à la lutte contre la brucellose bovine. Moniteur Belge, 1964.

Hadorn, D.C., Stark, K.D., 2008. Evaluation and optimization of surveillance systems for rare and emerging infectious diseases. Vet. Res. 39-57.

Knight-Jones, T.J., Hauser, R., Matthes, D., Stark, K.D., 2010. Evaluation of effectiveness and efficiency of wild bird surveillance for avian influenza. Vet. Res. 4, 41-50.

Martin, P.A., Cameron, A.R., Barfod, K., Sergeant, E.S., Greiner, M., 2007. Demonstrating freedom from disease using multiple complex data sources 2: case study–classical swine fever in Denmark. Prev. Vet. Med. 79, 98-115.

Martin, P.A., Cameron, A.R., Greiner, M., 2007. Demonstrating freedom from disease using multiple complex data sources 1: a new methodology based on scenario trees. Prev. Vet. Med. 79, 71-97.

Seleem, M.N., Boyle, S.M., Sriranganathan, N., 2010. Brucellosis: a re-emerging zoonosis. Vet. Microbiol. 140, 392-398.

Spencer, S., Sanna, M., Bakker, D., Pouillot, R., Garin-Bastuji, B., 2002. How to substantiate eradication of bovine brucellosis when aspecific serological reactions occur in the course of brucellosis testing. Vet. Microbiol. 90, 461-477.

Stark, K.D., 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 Serv Res, 6-20.

Welby, S., van den Berg, T., Marche, S., Houdart, P., Hooyberghs, J., Mintiens, K., 2010. Redesigning the serological surveillance program for notifiable avian influenza in Belgian professional poultry holdings. Avian Dis. 54, 597-605.

Wilkinson, L., 1993. Brucellosis. In The Cambridge World History of Human Disease. Edited by Kenneth F.K. Cambridge: University Press.

CHAPTER V: BOVINE TUBERCULOSIS SURVEILLANCE ALTERNATIVES IN BELGIUM

Prev Vet Med. 2012 Sep 15;106(2):152-61.

Bovine tuberculosis surveillance alternatives in Belgium.

Welby S, Govaerts M, Vanholme L, Hooyberghs J, Mennens K, Maes L, Van Der Stede Y.

Abstract

Belgium obtained the bovine tuberculosis (bTB) officially free status in 2003 (EC Decision 2003/467/EC). This study was carried out to evaluate the components of the current bTB surveillance program in Belgium and to determine the sensitivity of this program. Secondly, alternatives to optimize the bTB surveillance in accordance with European legislation (Council Directive 64/432/EEC) were evaluated.

Separate scenario trees were designed for each active surveillance component of the bTB surveillance program. Data from 2005-2009 regarding cattle population, movement and surveillance were collected to feed the stochastic scenario tree simulation model. A total of 7,403,826 cattle movement history records were obtained for the 2,678,020 cattle from 36,059 cattle herds still active in 2009. The current surveillance program sensitivity as well as the impact of alternative surveillance protocols was simulated in a stochastic model using 10,000 iterations per simulation.

The median (50% percentile) of the component sensitivities across 10,000 iterations was 0.83, 0.85, 0.99, 0.99 respectively for i) testing the cattle only during the winter screening ii) testing only imported cattle iii) testing only purchased cattle and iv) testing only all slaughtered cattle. The sensitivity analysis showed that the most influential input parameter explaining the variability around the output came from the uncertainty distribution around the sensitivity of the diagnostic tests used within the bTB surveillance. Providing all animals are inspected and post mortem inspection is highly sensitive, slaughterhouse surveillance was the most effective surveillance component. If these conditions were not met, the uncertainty around the mean sensitivity of this component was important. Using an antibody ELISA at purchase and an interferon gamma test during winter screening and at import would increase greatly the sensitivity and the confidence level of Belgium's freedom from bTB infection status.

INTRODUCTION

Although several European Union (EU) Member States have achieved the official bovine tuberculosis (bTB) free status (<0.1% annual herd prevalence), the possibility of a re-emergence of bTB infection cannot be excluded (EC Decision 2003/467/EC; EFSA, 2009).

Belgium has maintained the bTB officially free status since 2003 (EC Decision 2003/467/EC). Yet, sporadic outbreaks do still occur, as was recently the case in Germany and in the Netherlands (FASFC, 2010; Humblet et al., 2011b; Probst et al., 2010). Belgium had 7 reported breakdowns in 2003, 8 in 2004, 5 in 2005, 8 in 2006, 5 in 2007, 12 in 2008 and 2 in 2009. In 2010, no cases were detected (FASFC, 2010).

To date, despite the decreased prevalence following the sanitary measures implemented during the last century, Mycobacterium bovis still remains a constant (re-)emerging threat for animal and public health as well as for free animal trade in Europe and worldwide. Mycobacterium bovis, a slow-growing microaerobic bacterium, is the main causative agent of tuberculosis in cattle, and can infect a wide range of different species. If clinical signs appear, the main signs in cattle are wasting, weight loss, and fever. Because the infection does not always cause clinical signs, it might go unnoticed for several years before eventually being detected at slaughter by enlarged and/or caseating lymph nodes.

A national surveillance program according to the guidelines laid down in Council Directive 64/432/EEC and Royal Decree 17.10.2002 was in place in Belgium for many years. This program consists of four different surveillance components (SSC), each with its own sampling - and diagnostic process. The main aim of this ongoing surveillance program is to ensure the country is below the minimum required design prevalence required to be considered free from infection. In addition, the surveillance program should enable the early detection and eradication of sporadic cases. The Belgian Federal Agency for the Safety of the Food Chain (FASFC) commissioned this study to evaluate the sensitivity of the current surveillance for bTB in Belgium and how this program could be optimized to enable efficient detection of outbreaks and maintain the bTB officially free status. Several studies have investigated the efficacy of bTB surveillance components separately, in Belgium (Humblet et al., 2009a, 2011) and elsewhere (Corner et al., 1990; Frankena et al., 2007). In the present study the different bTB surveillance components were evaluated together and for their relative efficacy.

The scenario tree methodology developed by Martin et al. (2007a, b) has been proven to be a useful tool to quantify the sensitivity of a country's surveillance system (Frossling et al., 2009; Hadorn et al., 2002, 2008; Knight-Jones et al., 2010; Martin et

al., 2007a,b; More et al., 2011; Stark et al., 2006; Welby et al., 2009a,b, 2010). The surveillance system sensitivity is the probability of detecting an infection given the country is infected at the set design prevalence by testing all samples of the surveillance system. The sensitivity of the diagnostic test (TSe) is taken into account as well as the number of samples processed. Furthermore, the differential risk of infection for the specific risk groups in the cattle population are taken into account to quantify the benefits of targeting surveillance in those risk groups. However, when quantitative information about the key parameters affecting the differential risk of infection is not available, one needs to rely on assumptions to simulate values (Dohoo et al., 2009).

In the present study, the most representative values for the key parameters were estimated using historical data of the ongoing surveillance for bTB in Belgium. All available information regarding past outbreaks in Belgium was used to obtain quantitative values to feed the scenario tree model and obtain an output that reflected the Belgian situation, the sensitivity of its surveillance program and how it could be improved.

MATERIALS AND METHODS

Scenario tree methodology was chosen for the purpose of this study. Each event from infection to detection is represented by a node. A node has branches illustrating the possible event outcomes. Each outcome has a probability of occurrence. The combination of probabilities results in the probability of infection and detection for each limb, defined by a combination of nodes and branches, of the tree. The limbs define the different risk groups. The probabilities of detection are then combined to obtain the sensitivity of each SSC (CSe) and of the total surveillance system. Simulations are then carried out to see how changes in the current surveillance program could impact the CSe and surveillance system sensitivity, and how the surveillance program could be optimized in terms of efficiency of detection and provision of a guarantee of freedom from infection.

Belgian surveillance program

The four SSC of the current Belgian official surveillance program for bTB, in accordance with the guidelines laid down in Council Directive 64/432/EEC and the Royal Decree 17.10.2002, were:

- Imported cattle (IMP): all imported cattle are tested by single intradermal injection of bovine tuberculin (ID) at import. This SSC excludes young fattening calves (FC), which are sent to slaughter at the age of 6 months.
- Slaughtered cattle (SLGH): all slaughtered cattle undergo a post-mortem inspection (PM) at slaughterhouse.
- Purchased cattle (PUR): all purchased cattle, except FC for veal production, are tested with ID at purchase
- Winter screening (WS): ID is carried out on:
  - ✓ all cattle older than 6 weeks from herds considered as neighbour or contact herds of a suspected or confirmed bTB positive herd, after tracing-on and tracing-back investigation
  - ✓ all female cattle older than 24 months which belong to on-farm 'milk-selling' herds
  - ✓ all imported cattle, except FC for veal production, from non-bTB officially free Member States, are ID tested 3 consecutive years, during the winter period.

An animal that tested positive with ID is consecutively tested for confirmation by a comparative skin test, using bovine and avian tuberculin. Animals that test positive to this comparative ID will be slaughtered and all suspect lesions and/or lymph nodes at post-mortem examination will be tested by Ziehl-Nielsen staining and bacterial culture. In case of a positive result (isolation and identification of M. bovis) all animals of that herd will be slaughtered and a complete epidemiological investigation will be performed. A herd is considered free of bTB if all animals, older than 6 weeks, react negatively to the ID test 6 and 12 months after slaughtering all infected animals of a confirmed case herd, or 60 days following the introduction of (new) animals into a fully depopulated herd (FASFC, 2010; Royal Decree 17.10.2002).

A country or region is considered bTB officially free by the EU if for 6 consecutive years, 99.9% of the herds were tested and found negative. The annual herd prevalence must be <0.1% (Council Directive 64/432/EEC and Royal Decree 17.10.2002).

Scenario Tree

Four separate scenario trees, as described by Martin et al. (2007a), were designed for each SSC, namely, IMP, SLGH, PUR and WS. In each tree, the category nodes affecting the probabilities of infection were retained. Firstly considered was the risk category node 'importing from risk country'. This node was split into two branches (yes/no). This node was followed by the nodes 'previous bTB status' (positive/negative), 'animal movement rate' (low/high), 'herd type' (FC for veal production or other bovines), and 'herd size' (low, medium, high). Subsequently, these nodes led to different risk groups for which a herd effective probability of infection and an animal effective probability of infection were obtained based on the design prevalence at herd and within herd level (DPh, DPa) and the differential risk of infection of each node in each risk group. The category nodes affecting the probability of detection in each risk group for PUR, WS, IMP were the 'ID' test and for SLGH component, the 'PM' test. Figure 1 shows the general structure of the scenario tree. Each limb of the tree was defined by the combination of each category node branch probabilities. Only one possible combination is illustrated, but in the current study all possible combinations were accounted for.

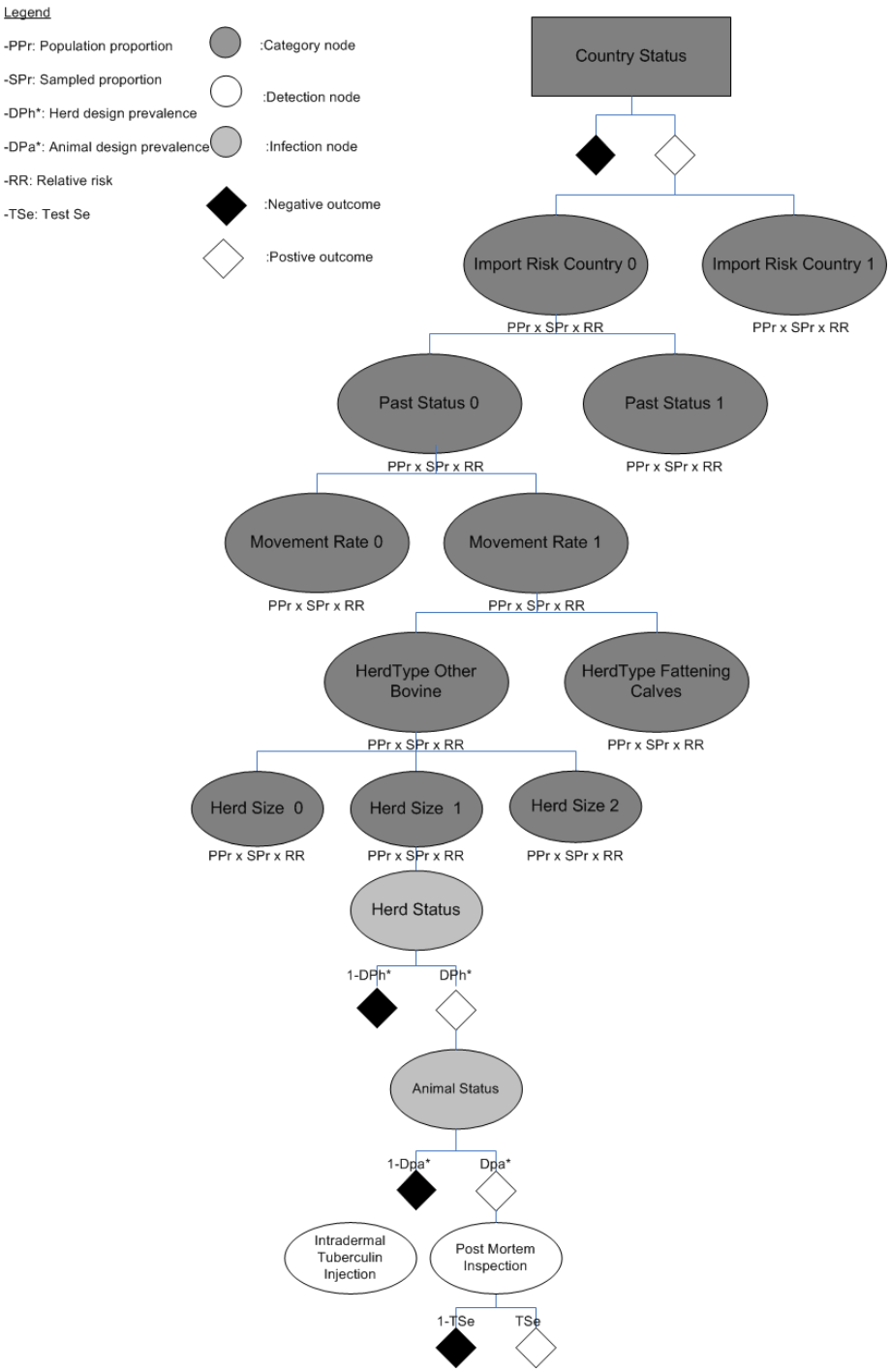


Figure 1: Scenario tree illustrating each step between the probability of infection and detection of bovine tuberculosis in Belgium. The Belgian cattle population is categorised in different nodes according to their differential risk of infection or detection. The different parameters at each node enabled to compute for each combination of nodes a sensitivity of detection at animal and herd level. Only one combination is illustrated here above for clarity purpose, but all combinations were considered in the study.

Node values

The choice and the sequence of the category nodes were developed following a review of literature (Humblet et al., 2009b, 2010) and consultation with four Belgian bTB experts from the national Veterinary Research Center (CODA-CERVA) and FASFC in Brussels. The Belgian national animal registration system (SANITEL) was then queried to extract data from the previous five years (2005-2009). The data were related to population data (cattle herd population structure), imported cattle, purchased cattle, slaughtered cattle, past status of bTB for each herd, and outbreak history of each reported bTB case in Belgium. The regional veterinary laboratories (ARSIA and DGZ) provided data regarding the cattle tested during the WS.

A total of 7,403,826 cattle movement history records were obtained for the 2,678,020 cattle and 36,059 herds still active in 2009. All datasets were standardized, formatted, merged and concatenated at herd level (SAS 9.2) and summarized in Table 1. For each herd that was still active in 2009 (reference year), the following parameters were obtained: import of cattle from risk country or not (binary variable), bTB status during the previous five years (yes/no), the average yearly movement rate (continuous variable), the herd type (FC or other bovines), herd size (continuous variable), the yearly number of slaughtered animals (continuous variable), number of purchased animals (continuous variable), number of imports (continuous variable), and the number of animals tested during WS (if WS was applicable, continuous variable).

Table 1: Number of herds and cattle tested in each surveillance component in 2009 in Belgium

Surveillance system component	Herds	Cattle
Slaughterhouse	26,138	769,016
Purchased	18,493	658,929
Imports	3,804	109,283
Winter screening	2,911	260,696

To categorize the reference population and the population sampled into the different category nodes branches, separate univariable analyses were conducted to determine cut-off values and the mean number of herds and animals in each node branch (SAS 9.2).

The differential risks of each nodes branch were obtained from a multivariable generalized linear model (SAS 9.2). The probability of having a bTB positive herd was modeled with a binomial distribution. The explanatory variables retained in the model, due to their biological plausibility or levels of significance, were: Import from risk countries (yes/no), previous status (positive/negative), movement rate (low/high), herd type (FC, other bovines) and herd size (small/medium/large). A pert distribution was fitted around the mean value of the relative risk estimate for each variable, using the 95% confidence intervals limits of the point estimate for the minimum and maximum value of the pert distribution. For the diagnostic test sensitivity values (TSe), literature (de la Rua-Domenech, 2006; Humblet et al., 2011; Cousins and Florisson, 2005) and expert opinion from the Veterinary and Agrochemical Research Centre (CODA-CERVA) and FASFC were consulted. The sources of information for the different nodes and branches input values are detailed in Table 2.

Table 2 Bovine tuberculosis (bTB) scenario tree nodes and branches cut-off values and sources of information.

Category nodes Infection	Branches	Cut off value, source of information	Input value
Import from risk country	Yes	Animal registration number	Uniform(1;1)
	No		Uniform(1;1)
Previous bTB status	Negative	Previous5 year TB history of the herd	Uniform(1;1)
	Positive		Uniform(30.7;30.7)
Movement rate	Low(≤2)	Univariate study 50th percentile	Uniform(1;1)
	High(>2)		Pert(2.7;4.9;10.5)
Herd type	Fattening Calves	Holding registration number	Uniform(1;1)
	Other bovine		Uniform(100;100)
Herd size	Small(≤10)	Univariate study 25thpercentile 75thpercentile	Uniform(1;1)
	Medium(>10≤25)		Uniform(1;1)
	Large(>25)		Pert(0.09;0.32;1.05)
Category nodes Detection			
Intradermal skin test	Positive	Literature review	Pert(0.54;0.68;0.95)
	Negative	Expert opinion	Uniform(1;1)
Post mortem inspection	Positive	Literature review	Pert(0.5; 0.7; 0.99)
	Negative	Expert opinion	Uniform(1;1)

Model description

Spreadsheets were created in Excel 2007 to represent each surveillance component (IMP, PUR, SLGH, WS). Distributions were fitted to each input variable taking into account the variability as well as the uncertainty of the key parameters (ModelRisk 3.0).

Adjusted relative risks of infection

The relative risk estimates (RR) obtained from above, used to describe the herd effective probability of infection (EPIH) in each risk group (i), were adjusted (AR) to take in account the herd population structure (PPr) (Eq. (1, 2)). The herd effective probability of infection for each risk group was then obtained from the ARi and the DPh (0.1%), obtained from Council Directive 64/432/EEC.

$$AR_i = \frac{RR_i}{\sum_{i=1}^{n} RR_i * PPr_i} \text{(Eq.1)}$$

$$EPIH_i = AR_{IR_i} * AR_{S_i} * AR_{MR_i} * AR_{HT_i} * AR_{HS_i} * DPh \text{(Eq. 2)}$$

Where IR_i, S_i, MR_i, HT_i, HS_i stands for the nodes import from risk country, past bTB status, movement rate, herd type and herd size respectively.

Animal and herd sensitivity estimation (ASe, HSe)

The combination of TSe and sampled population proportion (SPr) enabled to obtain the effective probability of detection at animal level (EPDA) for each risk group (i) (Eq. (3)).

$$EPDA_i = SPr_{IR_i} * SPr_{S_i} * SPr_{MR_i} * SPr_{HT_i} * SPr_{HS_i} * TSe \text{ (Eq. 3)}$$

The sensitivity at animal level (ASe) and the herd level sensitivity (HSe) for each risk group (i) can then be estimated by the following 2 equations (Eq. (4, 5)),

$$ASe_i = 1 - (1 - (EPDA_i * \frac{n_{a_i}}{N_{a_i}}))^{N_{a_i} * DPa} \text{(Eq. 4)}$$

$$HSe_i = 1 - (1 - (ASe_i * \frac{n_{h_i}}{N_{h_i}}))^{N_{h_i} * EPIH_i} \text{ (Eq. 5)}$$

where $\frac{n_{a_i}}{N_{a_i}}$ is the average number of animals tested, divided by the average numbers of animals per herd for each risk and $\frac{n_{h_i}}{N_{h_i}}$ represents the average number of herds sampled divided by the total herd population. No values regarding the minimum DPa were available in the legislation, therefore these values were estimated from previous bTB outbreaks (2005-2009). The best distribution value for this parameter was selected based on the akaike information criterion (AIC) (Normal (0.17;0.20)).

Component and system sensitivity estimation

The component sensitivity (CSe) for each component investigated was estimated from the product of the individual herd's sensitivity for each limb of the tree (Eq. (6)):

$$CSe_i = 1 - \prod(1 - HSe_i) \text{ (Eq. 6)}$$

Once all these component sensitivities were estimated, the surveillance system sensitivity (SSe) could be computed by (Eq. (7)):

$$SSe_i = 1 - \prod(1 - CSe_i) \text{ (Eq. 7)}$$

Model simulations

The following scenarios were simulated through a stochastic iteration process (10,000 Iterations/Simulation) (ModelRisk 3.0):

The CSe of each current ongoing SSC was estimated

The sample size and PM TSe were investigated to determine what measures could affect the SLGH CSe

The test sensitivity of the ID for WS and PUR was investigated to determine what reflected most accurately the TSe given the circumstances in which the ID is conducted.

Other diagnostic assays were simulated to measure the impact on CSe for WS, PUR and IMP.

The impact of targeting in certain risk groups was simulated as well for PUR, SLGH and WS.

Finally, the confidence level around our posterior probability of freedom from infection was then simulated from year to year (2005-2009) given the discovery of new outbreaks (Eq. (8, 9)).

$$\mathrm{Pfree}_{t_i} = \frac{\text{1-Pinf}_{ti\text{-}1}}{\text{1-Pinf}_{t_i\text{-}1} * \mathrm{CSe}_{t_i\text{-}1}} \quad \text{(Eq. 8)}$$

$$Pinf_{t_i} = \left(1 - Pfree_{t_i-1}\right) + Pintro_{t_i} - \left(Pintro_{t_i} * \left(1 - Pfree_{t_i-1}\right)\right) \quad \text{(Eq. 9)}$$

The posterior probability of freedom (Pfree) is obtained given prior probability of infection (Pinf) of the previous year and the CSe of the previous year. The Pinf was set to 0.001, the minimum design prevalence, the first year (<2005). Then, for the following years, it was equal to $1 - \mathrm{Pfree}_{t_i-1}$. The probability of introduction (Pintro) was equal to the bTB annual prevalence.

Model validation

An internal validation was done by running a sensitivity analysis to determine which risk groups of each SSC were most influential on the CSe.

An external validation was done in parallel to validate the output of this model. For this purpose a generalized estimating equation (GEE) model was built to investigate what was the most significant reason that led to the detection of outbreak herds over the past five years, based on the p-value (p-value < 0.05) of the Wald test statistic (SAS 9.2).

RESULTS

The median (50% percentile) of the component sensitivities across 10,000 iterations was 0.83, 0.85, 0.99, 0.99, respectively, for i) testing the animals only during the winter screening ii) testing only imported cattle iii) testing only purchased cattle and iv) testing only all slaughtered cattle. The uncertainty distribution of the CSe was wide around IMP SSC (Fig. 2). In 75% of the 10,000 iterations during the PUR and SLGH simulations, the CSe simulations remained above 0.9 (Fig. 2).

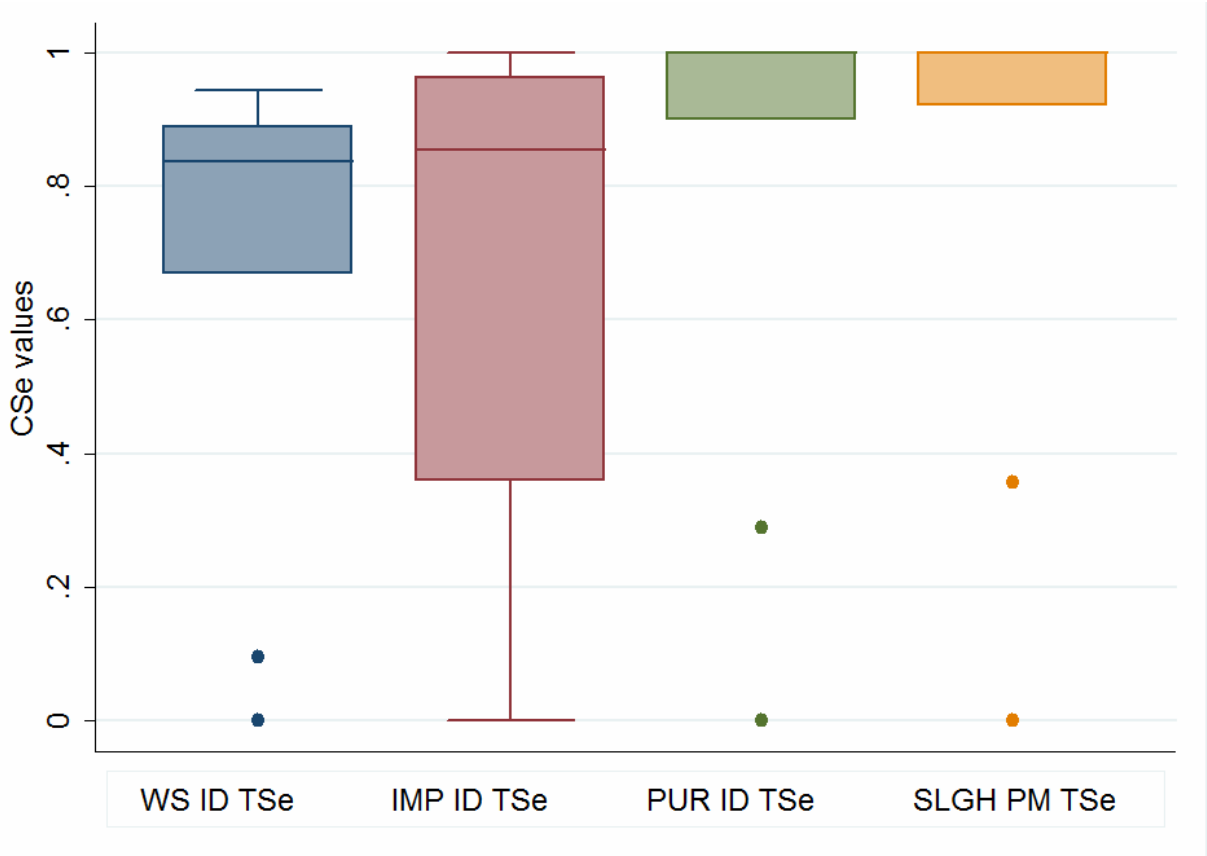


Figure 2: Simulations history of each surveillance system component sensitivity (CSe) for bovine tuberculosis detection in Belgium. Box plots representing the 25th (bottom of the box) and 75th (top of the box) percentiles of the iterations history for each simulated scenario (CSe (Y axis) (winter screening (WS) with intradermal skin test (ID) (Test sensitivity (TSe) (defined by a pert distribution (Pert))=Pert(0.54; 0.68; 0.95)), import (IMP) with ID test (TSe =Pert(0.54; 0.68; 0.95)), purchase (PUR) with ID test (TSe =Pert(0.54; 0.68; 0.95)), slaughter (SLGH) with the post mortem inspection (PM) (TSe =Pert(0.5; 0.7; 0.99)) and DPa=Normal(0.17; 0.2)) (X axis). The midline corresponds to the 50th percentile (median).The lower and upper whiskers represent the 5th and 95th percentile. The height of the box shows the skew of the

data around the median. For PUR and SLGH, the 50th percentile was the same value as the 75th percentile. The dots represent the outliers.

To determine what had the higher impact in SLGH SSC, simulations were carried out. Fixing the TSe of SLGH at 0.99, 0.75, or 0.5 did not have as much impact as reducing the sample size. When 50% of the sample size was simulated, the CSe showed a very wide uncertainty range. In all simulations, however, the mean CSe remained above 0.9 (Fig. 3).

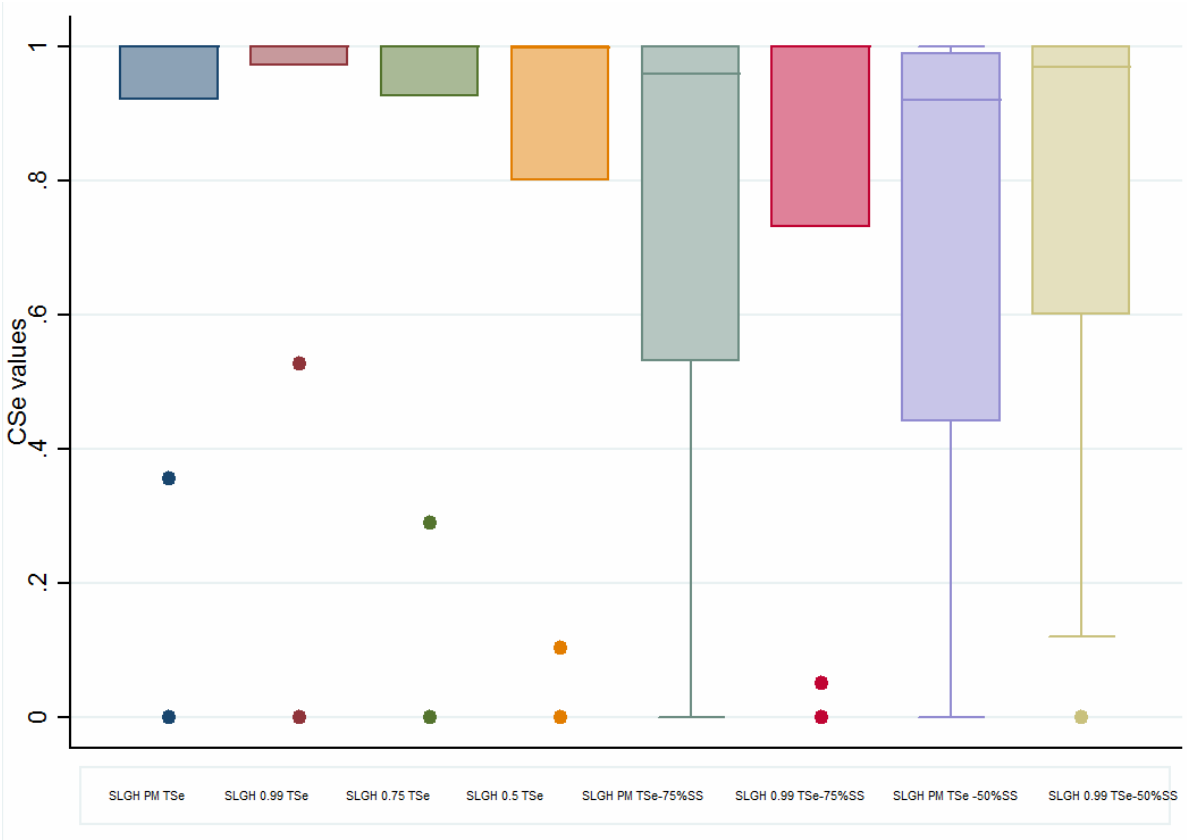


Figure 3: Simulations of the slaughter (SLGH) surveillance component sensitivity (CSe) for bovine tuberculosis detection in Belgium and some scenario simulations. Box plots representing the 25th (bottom of the box) and 75th (top of the box) percentiles of the iterations history for each simulated scenario (CSe (Y axis) (SLGH with post mortem inspection (PM) (Test sensitivity (TSe) (defined by a pert distribution (Pert))=Pert(0.5; 0.7; 0.99)), SLGH with PM (TSe= 0.99)), SLGH with PM (TSe=0.75)), SLGH with PM (TSe= 0.5), SLGH with PM (TSe=PERT(0.5; 0.7; 0.99)) and 75% cattle inspected), SLGH with PM (TSe= 0.99) and 75% cattle inspected, SLGH with PM (TSe=Pert(0.5; 0.7; 0.99)) and 50% cattle inspected, SLGH with PM (TSe= 0.99)and 50% cattle inspected), and design prevalence at animal level

(defined by a normal distribution (Normal))=Normal(0.17; 0.2)) (X axis). The midline corresponds to the 50th percentile. The lower and upper whiskers represent the 5th and 95th percentile. The height of the box shows the skew of the data around the median. The dots represent the outliers.

History of the outbreaks over the previous five years indicated that WS was more effective in detecting outbreaks then PUR. Different ID TSe values were simulated to reflect the difference of TSe depending on the circumstances in which this test is carried out (Fig. 4).

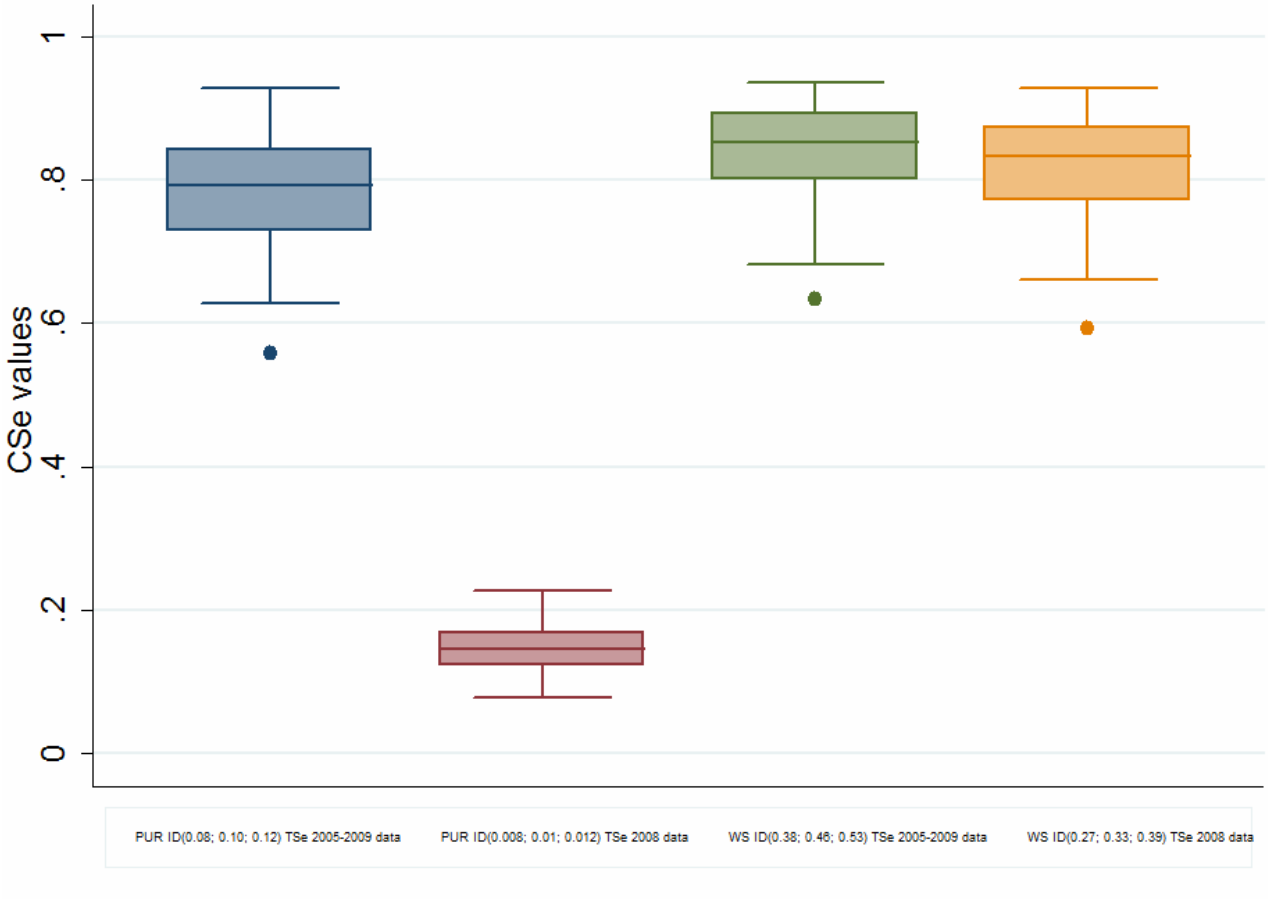


Figure 4: Simulations of the winter screening (WS) and purchase (PUR) surveillance component sensitivity (CSe) for the detection of bovine tuberculosis in Belgium and some scenario simulations. Box plots representing the 25th (bottom of the box) and 75th (top of the box) percentiles of the iterations history for each simulated scenario (CSe (Y axis) for each surveillance component (PUR with intradermal skin test (ID) (Test sensitivity (TSe) (defined by a pert distribution (Pert))=Pert(0.08; 0.10; 0.12)) obtained from 2005-2009 data, PUR with ID (TSe=Pert(0.008; 0.010; 0.012)), WS with ID (TSe=Pert(0.38; 0.46; 0.53)) obtained from 2005-2009 data, WS with ID

(TSe=Pert(0.27; 0.33; 0.39), and design prevalence at animal level (defined by a normal distribution (Normal))=Normal(0.17;0. 20)) (X axis). The midline corresponds to the 50th percentile. The lower and upper whiskers represent the 5th and 95th percentile. The height of the box shows the skew of the data around the median. The dots represent the outliers.

We then investigated how the CSe could be improved in each SSC (Fig.5). Antibody ELISA (Ab ELISA) was considered for the PUR component; this increased the CSe and greatly reduced the uncertainty around the CSe. Interferon gamma (INFγ) was considered for WS and IMP. Only sampling FC in SLGH and PUR, and outbreak herds with a INFγ in WS was also simulated. This targeted approach was of limited value.

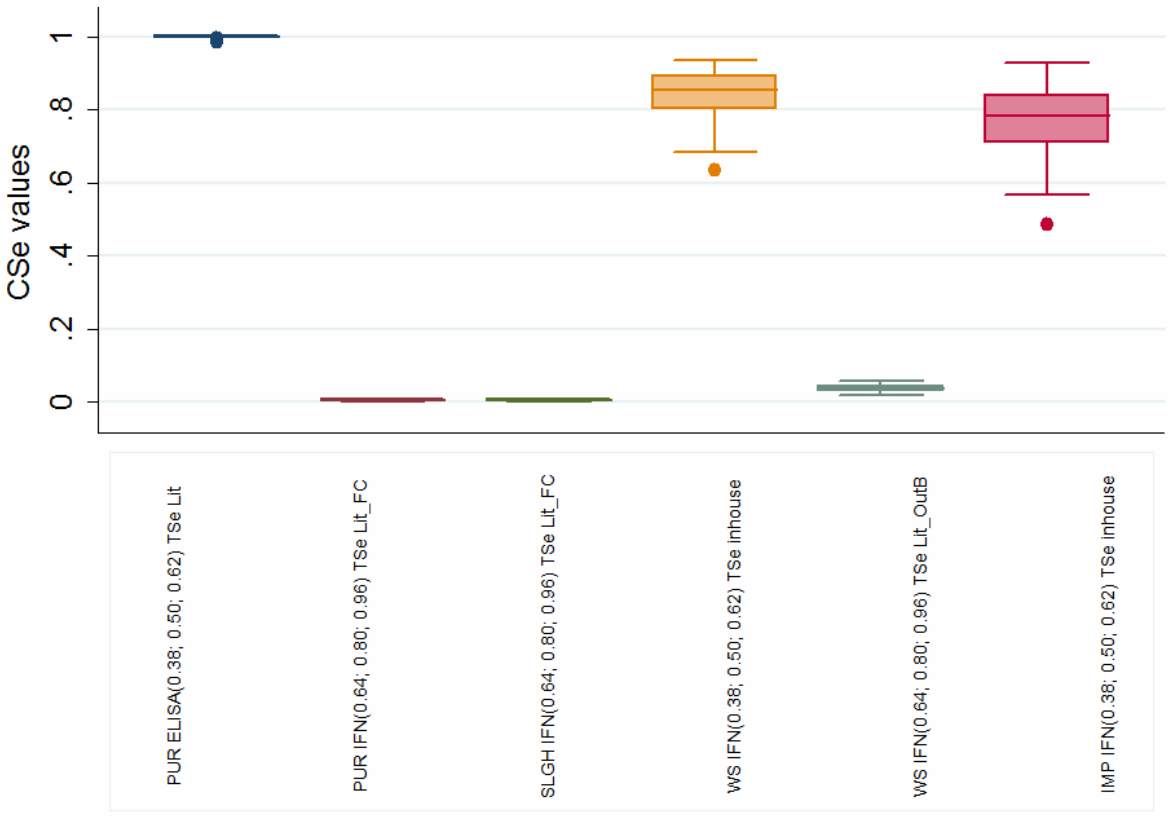


Figure 5: Simulations of alternative options regarding diagnostic assays for each bovine tuberculosis surveillance component in Belgium. Box plots representing the 25th (bottom of the box) and 75th (top of the box) percentiles of the iterations history for each simulated scenario (Component sensitivity (CSe) (Y axis) for each surveillance component (purchase (PUR) with Ab ELISA (Test sensitivity (TSe)=literature value), PUR targeting only fattening calves (FC) with IFNγ

(TSe=literature value), slaughter targeting only FC with IFNγ (TSe=literature value), winter screening (WS) with IFNγ (TSe=CODA-CERVA in house validation value), WS targeting only past outbreaks with IFNγ (TSe=literature value), import (IMP) with IFNγ (TSe=literature value), and design prevalence at animal level (defined by a normal distribution (NORMAL))=NORMAL(0.17;0.02)). The midline corresponds to the 50th percentile. The lower and upper whiskers represent the 5th and 95th percentile. The height of the box shows the skew of the data around the median. The dots represent the outliers.

The TSe values of Ab ELISA (0.38; 0.5; 062), and INFγ (0.64; 0.80; 0.96) used for targeting were those found in literature (van Asseldonk et al., 2005). The INFγ (0.38; 0.5; 062) values used for the whole population sampled during WS and IMP were the values found in the in-house validation test of the INFγ test (Govaerts et al., 2009).

The posterior probability of infection was modeled for the past five years (Fig.6).The probability of infection was higher in years where more outbreaks were detected (2008), but always remained below the minimum design prevalence prescribed by the legislation. The PUR and SLGH SSCs each separately provide a good guarantee for Pfree, as seen for the surveillance system sensitivity. In contrast, testing either IMP or WS were less efficient.

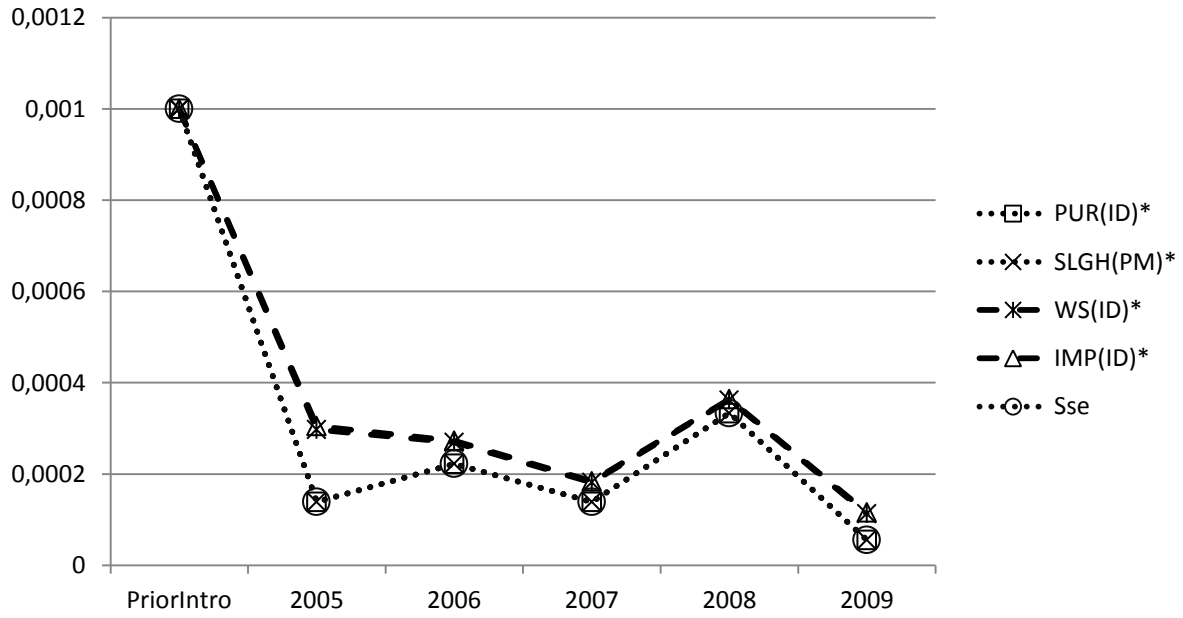


Figure 6: Posterior probability of infection of bovine tuberculosis in Belgium using each surveillance component (purchase (PUR), slaughterhouse (SLGH), winter screening (WS), import (IMP) with intradermal skin test (ID) or post mortem inspection (PM) and total surveillance system sensitivity (SSe))

From the internal validation, the sensitivity analysis of this model showed that the diagnostic method (TSe) was the most determinant factor on the variation of the output, followed by the intra-herd prevalence. Large herds of adult bovines, where movement rate was high, where no bTB infections were observed previously, and where no imports from risk countries were registered, were the most influential on the CSe in each SSC. Thus targeting surveillance or continuing the ongoing surveillance in those risk groups is the best approach to increase CSe of each SSC.

The GEE model, carried out for external validation, showed that the most efficient surveillance method over the previous five years was SLGH (1.69, p-value ≤ 0.01), followed by WS (1.34, p-value ≤ 0.0001) for the detection of outbreaks. The PUR surveillance component was non-significant (p-value=0.91).

DISCUSSION

The output of this study has underlined features of the current surveillance program in Belgium, such as the importance of slaughterhouse surveillance (SLGH). It can be concluded from the simulations and the results of the external validation model, that SLGH provides the best CSe in Belgium. One possible explanation for the high sensitivity in the slaughterhouse component is the large sampling coverage. However, it is obvious that the efficiency of this component is highly dependent on the sensitivity of the post mortem examination (during meat inspection by veterinarians) and on the sample size. This was demonstrated by the simulation exercise and the sensitivity analysis (Fig. 3). An important feature found during this simulation exercise is that in 25 % of the iterations the sensitivity of the SLGH component was below 0.92 (Fig.2). This means that considering the different input parameters (TSe, population size, sample size and relative risk estimates), an infection present in the country at 0.1 % design prevalence could go undetected if only slaughterhouse surveillance were carried out. It is the combination of all surveillance components that counter this lack of CSe seen in SLGH. High variability in the sensitivity of visual inspection is also mentioned in literature (Corner et al., 1990; Cousin and Florisson, 2005; Frankena et al., 2011; More and Good, 2006; Ryan et al., 2006). In the past, slaughterhouse surveillance has proven to be effective at national level, as in some other Member States, for the detection of bTB

outbreaks that were not detected at purchase by the means of ID testing (Collins, 2006; Humblet et al., 2011b; More and Good, 2006; Probst et al., 2010; van Asseldonk et al., 2005). The GEE model, developed for validation purpose, supports these observations.

Although the WS had low CSe, partly due the small fraction of the population sampled within this component, it was seen that WS was actually more efficient than PUR when the TSe of ID test was corrected, as shown in Fig. 4. The targeted approach of this WS component probably explains this high CSe, as shown by the GEE validation model. Data from the past five years showed that WS was more effective at detecting outbreaks than the PUR component in Belgium. Although the same ID skin test was applied, the TSe varied given the situations. The TSe of the WS was assigned a higher value because it was expected that in circumstances of outbreak investigation, the vigilance of the veterinarians carrying out ID would be higher compared to that in routine testing in normal circumstances, e.g. testing at purchase. When new simulations were carried out taking into account the difference of TSe between PUR and WS, WS turned out to have the highest CSe value (Fig. 4). However in 25% of the iteration process history, the CSe of WS was only 0.75. To counter the lack of sensitivity observed with the ID test alone and because the proportion of the population investigated during WS is limited, implementing an IFN-γ during WS in parallel of the ID test, could be a good, cost-effective alternative. This has been seen in the study and supported by literature (Collins et al., 2006; More and Good, 2006).

The PUR SSC in the current bTB program in Belgium was of limited value. This reflects the Belgian experience, where for several years no outbreaks were detected during PUR. However, testing purchased animals remains important because cattle movement constitutes one of the major routes of bTB spread. Implementing a different diagnostic test, such as the Ab ELISA, could be a good cost effective alternative to counter this lack of Se of the ID carried out during PUR, as seen in Fig. 5. But further studies would be needed to confirm the TSe value of the Ab ELISA simulated in this model. Furthermore, keeping good records of all animal movements is crucial to enable efficient tracing-back and tracing-on of infected animals or herds. This traceability is the only way to detect the route of dissemination of a bTB infection

(Collins, 2006; Gilbert et al., 2005; Humblet et al., 2011; More et al., 2011; Probst et al., 2010).

The IMP SSC had a wide range of uncertainty around the CSe. During the previous five year outbreak history in Belgium, no case herd was found from herds that imported from risk countries. So targeting that risk group did not increase the probability of freedom from infection in our simulations. The low prevalence of infection within the country probably constitutes the source of these sporadic outbreaks.

Targeting the surveillance on veal calves did not seem to enhance the CSe of each SSC (Fig.5). This is mainly due to the fact that in this study FC was not considered to be at risk of being infected by bTB, following the risk factor analysis. To date, no outbreaks have been found in FC in Belgium. This is not the case in some neighboring countries. At the moment there is no target sampling planned in the FC population by the actual surveillance program.

A herd level spatial cluster analysis was performed to investigate if any significant spatial clustering of outbreaks existed at municipality level in (SaTScanTM 2005). A Poisson distribution was chosen to model the differential risk of municipalities containing positive herds. Two significant clusters were found: one in the province of Antwerp (relative risk of 19 with p-value of 0.001) and one cluster covering the provinces Liege, Limburg and Flemish Brabant (relative risk 8.5, p-value 0.033). However, it was decided not to take into account spatial aspects in the final scenario tree, as WS already accounts for this targeted approach. Walravens et al. (2006) have shown that bTB strain could still re-emerge five years after the whole herd was slaughtered and replaced. This could suggest the presence of an environmental or wildlife reservoir. To date, no bTB cases have been reported in wildlife. However, no active surveillance has been carried out and the conclusion is based only on passive reporting. The possibility of a wildlife reservoir should not be ruled out based on experiences in some neighboring Member States (Abernethy et al., 2006; Anses, 2011; More et al., 2011; More and Good, 2006). Implementing wildlife active surveillance around these clusters should be considered to rule out this possibility. The spatial clusters, found in this study, were all near the Belgian national borders (provinces of Antwerp, Limbourg, Liege and Hainaut) suggesting that these regions are at high risk. However, it might be as well that the level of disease awareness was

higher in those areas, thus enhancing the number of cases found. Similar findings were observed in other publications (Collins, 2006; Elbers et al., 2010; Humblet et al., 2011).

CONCLUSION

Depending on the infection status of a country, two situations can be distinguished. In an endemic situation, trade of animals should be controlled because this event can constitute a high risk of dissemination. In an infection-free situation, testing and controlling all possible routes of introduction is a key element of the surveillance program. Eradication success depends on a program that enables early detection of the infection. From the simulation results of this study, it was found that individual inspection of carcasses at slaughterhouse (meat inspection) plays a key role in the surveillance of bTB but mainly depends on the sensitivity of the meat inspection, which is the probability that veterinarians detect bTB lesions during meat inspection. The ID test carried out at PUR and WS was of limited value. To increase the sensitivity at PUR and WS, an Ab ELISA could be used during PUR and an IFN-γ during WS, as simulated in this study and published elsewhere (Collins, 2006; More and Good, 2006; Ramirez-Villaescusaet al., 2010; van Asseldonk et al., 2005; Vanholme, 2009). However, further studies are recommended to confirm these hypotheses.

Obtaining empirical figures for the test sensitivities would enable us to increase the validity of this model. Nevertheless, the simulations carried out in this study provided good insight into the key elements to be considered when implementing or revising the bTB surveillance program in Belgium such as required by European legislation and International standards.

REFERENCES

Anses (Agence Natioanle de Securite Sanitaire de L'Alimentation, de L'Environnement et du Travail), 2011. Tuberculose et faune sauvage. Catherine Collignon (Ed.), Maisons-Alfort Cedex, France.

Abernethy DA, Denny, GO Menzies, FD, Mc Guckian P, Honhold N, Roberts AR, 2006. The Northern Ireland programme for the control and eradication of Mycobacterium bovis. Vet Microbiol. 112: 231-237.

Cousins DV, Florisson N, 2005. A review of tests available for use in the diagnosis of tuberculosis in non-bovine species. Rev Sci Tech. 24: 1039-1059.

Collins JD, 2006. Tuberculosis in cattle: strategic planning for the future. Vet Microbiol. 112: 369-381.

Corner, LA, Melville, L, Mc Cubbin, K, Small, KJ, Mc Cormick, BS, Wood, PR, Rothel, J.S., 1990. Efficiency of inspection procedures for the detection of tuberculous lesions in cattle. Aust Vet J. 67: 389-392.

de la Rua-Domenech R, Goodchild AT, Vordermeier HM, Hewinson RG, Christiansen KH, Clifton-Hadley RS, 2006. Ante mortem diagnosis of tuberculosis in cattle: a review of the tuberculin tests, gamma-interferon assay and other ancillary diagnostic techniques. Res Vet Sci. 81: 190-210.

Dohoo, I, Martin, W Stryhn, H (Eds.), 2009. Veterinary Epidemiologic Research. VER Inc, Canada.

European Commission (EC), 2003. Commission Decision 2003/467/EC establishing the official tuberculosis, brucellosis, and enzootic-bovine-leukosis-free status of certain Member States and regions of Member States as regards bovine herds. Official Journal L156, 74-78.

European Economic Community (EEC), 1964. Council Directive 64/432/EEC on animal health problems affecting intra-Community trade in bovine animals and swine. Official Journal 121, 1977 – 2012.

European Food Safety Authority (EFSA), 2009. The Community Summary Report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in the European Union in 2008. EFSA Journal.

Elbers, ARW, Gorgievski-Duijvesteijn, MJ, Zarafshani, K, Koch, G, 2010. To report or not to report: a psychosocial investigation aimed at improving early detection of avian influenza outbreaks. Rev. Sci. Tech. Off. Int. Epiz. 29 (3), 435-449.

Federal Agency for the Security of the Food Chain (FASFC), 2010. Tuberculosis, In: Trends and sources report on zoonotic agents in Belgium. Gil Houins (Ed.), Belgium, Brussels, pp. 89-92.

Frankena K, White PW, O'Keeffe O, Costello E, Martin SW, Grevenhof EM van, More SJ, 2007. Quantification of the relative efficiency of factory surveillance in the disclosure of tuberculosis lesions in attested Irish cattle. Vet Rec. 161: 679-684.

Frossling J, Agren EC, Eliasson-Selling L, Lewerin SS, 2009. Probability of freedom from disease after the first detection and eradication of PRRS in Sweden: scenario-tree modelling of the surveillance system. Prev Vet Med. 91: 137-145.

Gilbert M, Mitchel LA, Bourn D, Mawdsley J, Clifton-Hadley R, Wint W, 2005. Cattle movements and bovine tuberculosis in Great Britain. Nature 435: 491-496.

Govaerts M, Wolucka B, Murias J, Lammerty J., 2009. Convention RT 07/6 TUB : Développement et validation de méthodes alternatives pour le dépistage de masse de la tuberculose bovine (TUB) 01 VI 2008 – 31 V 3011. Rapport Intermédiaire d'Activités, 1ère année.

Hadorn DC, Rufenacht J, Hauser R, Stark KD, 2002. Risk-based design of repeated surveys for the documentation of freedom from non-highly contagious diseases. Prev Vet Med. 56: 179-192.

Hadorn DC, Stark KD, 2008. Evaluation and optimization of surveillance systems for rare and emerging infectious diseases. Vet Res. 39: 57.

Humblet MF, Walravens K, Salandre O, Boschiroli ML, Gilbert M, Berkvens D, Fauville-Dufaux M, Godfroid J, Dufey J, Raskin A, Vanholme L, Saegerman C, 2009a. Première évaluation et validation des pratiques d'intradermo- tuberculination des vétérinaires praticiens ruraux fondées sur un questionnaire d'enquête épidémiologique anonyme. Epidémiol et Santé Anim. 55 : 7-15.

Humblet MF, Boschiroli ML, Saegerman C, 2009b. Classification of worldwide bovine tuberculosis risk factors in cattle: a stratified approach. Vet Res. 5: 40-50.

Humblet MF, Gilbert M, Govaerts M, Fauville-Dufaux M, Walravens K, Saegerman C, 2010. New assessment of bovine tuberculosis risk factors in Belgium based on nationwide molecular epidemiology. J Clin Microbiol. 48: 2802-2808.

Humblet MF, Moyen JL, Bardoux P, Boschiroli ML, Saegerman C, 2011. The Importance of Awareness for Veterinarians Involved in Cattle Tuberculosis Skin Testing. Transbound Emerg Dis., 58: no. doi: 10.1111/j.1865-1682.2011.01228.x

Knight-Jones TJ, Hauser R, Matthes D Stark KD, 2010. Evaluation of effectiveness and efficiency of wild bird surveillance for avian influenza. Vet Res. 4: 41-50.

Martin PA, Cameron AR and Greiner M, 2007a. Demonstrating freedom from disease using multiple complex data sources 1: a new methodology based on scenario trees. Prev Vet Med., 79: 71-97.

Martin PA, Cameron AR, Barfod K, Sergeant ES, Greiner M, 2007b. Demonstrating freedom from disease using multiple complex data sources 2: case study--classical swine fever in Denmark. Prev Vet Med. 79 : 98-115.

MB (Moniteur Belge), 2010. Arrêté royal modifiant l'arrêté royal du 17 octobre 2002 relatif à la lutte contre la tuberculose bovine. Moniteur Belge, 2003022000.

More SJ, Good M, 2006. The tuberculosis eradication programme in Ireland: a review of scientific and policy advances since 1988. Vet Microbiol. 112: 239-251.

More SJ, Cameron AR, Greiner M, Clifton-Hadley RS, Rodeia SC, Bakker D, Salman MD, Sharp JM, De Massis F, Aranaz A, Boniotti MB, Gaffuri A, Have P, Verloo D, Woodford M, Wierup M, 2009. Defining output-based standards to achieve and maintain tuberculosis freedom in farmed deer, with reference to member states of the European Union. Prev Vet Med. 90(3-4): 254-67.

Probst C, Freuling C, Moser I, Geue L, Kohler H, Conraths FJ, Hotzel H, Liebler-Tenorio EM, Kramer M, 2010. Bovine tuberculosis: making a case for effective surveillance. Epidemiol Infect. 1: 105-112.

Ramirez-Villaescusa AM, Medley GF, Mason S, Green LE, 2010. Risk factors for herd breakdown with bovine tuberculosis in 148 cattle herds in the south west of England. Prev Vet Med. 95: 224-230.

Ryan TJ, Livingstone PG, Ramsey DS, de Lisle GW, Nugent G, Collins DM, Buddle BM, 2006. Advances in understanding disease epidemiology and implications for control and eradication of tuberculosis in livestock: the experience from New Zealand. Vet Microbiol. 112: 211-119.

Stark KD, Regula G, Hernandez J, Knopf L, Fuchs K, Morris RS, Davies P, 2006. Concepts for risk-based surveillance in the field of veterinary medicine and veterinary public health: review of current approaches. BMC Health Serv Res. 6: 20.

van Asseldonk MA, van Roermund HJ, Fischer EA, de Jong MC, Huirne RB, 2005. Stochastic efficiency analysis of bovine tuberculosis-surveillance programs in the Netherlands. Prev Vet Med. 69: 39-52.

Walravens K, Allix C, Supply P, Rigouts L, Godfroid J, Govaerts M, Portaels F, Dufey J, Vanholme L, Fauville-Dufaux M, Saegerman C, 2006. Dix années d'épidémiologie

moléculaire de la tuberculose bovine en Belgique. Epidémiol et Santé Anim. 49 : 103-111.

Welby S, Fretin D, Hooyberghs J, Vanholme L, Godefroid J, Van der Stede Y, 2009a. Evaluation du programme de surveillance pour la brucellose et la leucose bovine enzootique en Belgique. , Epidémiol et Santé Anim. 55 : 1-6. .

Welby S, Meroc E, Faes C, Mintiens K, De Clercq K, Hooyberghs J, Van der Stede Y, 2009b. Demonstrating the absence of certain BT serotypes: Evaluation of the different surveillance components prescribed by Regulation (EC 1266/2007). EPIZONE Report WP 6.6.

Welby S, van den Berg T, Marche S, Houdart P, Hooyberghs J, Mintiens K, 2010. Redesigning the serological surveillance program for notifiable avian influenza in Belgian professional poultry holdings. Avian Dis. 54: 597-605.

CHAPTER VI: BLUETONGUE SURVEILLANCE SYSTEM IN BELGIUM: A STOCHASTIC EVALUATION OF ITS RISK-BASED APPROACH EFFECTIVENESS

Prev Vet Med. 2013 112(1-2):48-57.

Bluetongue surveillance system in Belgium: a stochastic evaluation of its risk-based approach effectiveness.

Welby S, Meroc E, Faes C, De Clercq K, Hooyberghs J, Mintiens K, Van der Stede Y.

Abstract

Background: The aim of this study was to assess the sensitivity of the four major bluetongue surveillance components implemented in Belgium in 2007 for farmed animals and prescribed by the European Union regulation ; winter serological screening, sentinel system, passive clinical surveillance, export testing. Scenario tree methodology was used to evaluate the relative sensitivity of detection and targeted approach of each component in terms of early detection and freedom of infection substantiation. Field data collected from the previous year's outbreaks in Belgium were used to determine the risk groups to be considered. Results: The best sensitivities at herd level, taking into account the diagnostic test sensitivity, design prevalence and the number of animals tested within a herd were obtained with the winter screening and sentinel component. The sensitivities at risk group level, taking into account the obtained herd sensitivity, effective probabilities of infection and number of herds tested were high in all components, except for the export component. Component sensitivities ranged between 0.77 and 1 for all components except for the export component with a mean value of 0.22 (0.17-0.26). In terms of early detection, the probability of detection was best using the passive clinical component or the sentinel component. Sensitivity analysis showed that the passive clinical component sensitivity was mostly affected by the diagnostic process and the number of herds sampled. The sentinel and export components sensitivity were mainly affected by the relative risk estimates whereas the winter screening component was mainly affected by the assumptions about the design prevalence. Conclusions: This study revealed interesting features regarding the sensitivity of detection and early detection of different components and their risk based approach as requested by the international standards.

INTRODUCTION

Bluetongue is an arthropod-borne viral disease affecting both wild and domestic ruminants. The distribution of the bluetongue virus (BTV) is therefore limited to those regions where competent vector species are present and its transmission to those times of the year when climatic conditions are favourable for the cycle of transmission (Mellor et al., 2008; Wilson and Mellor, 2009). The vast majority of BTV infections are

subclinical. Though cattle are more likely to be infected by a biting midge, clinical signs are more apparent in sheep (Elbers et al, 2008a, b; Saegerman et al., 2010). BTV can cause mild to spectacular outbreaks and has an adverse impact on worldwide trade. Thus, it appears on the list of diseases notifiable to the World Organization for Animal Health (OIE).

Following the epidemics of BTV-8 in 2006, 2007 and 2008, Commission Regulation (EC) No 1266/2007 (EC, 2007) last amended in May 2012, prescribed the implementation of mandatory surveillance systems composed of i) passive clinical surveillance, ii) sentinel surveillance, iii) a combination of serological and/or virological surveillance, as well as targeted risk-based monitoring. More and more, rather than prescribing fixed guidelines, the aim of these current regulations are oriented towards minimum requirements to be fulfilled. As a consequence, regulations are more flexible and allow each member state to adapt its own surveillance activities and prove the effectiveness of its system towards the required objectives. The present study was carried out in this context. The four major BTV surveillance components implemented in Belgium in 2007 (winter screening, sentinel, export, and passive clinical) were evaluated for the relative efficacy in terms of their risk-based design, early detection and substantiation of infection freedom.

The scenario tree methodology, developed by Martin et al. (Martin et al., 2007a, 2007b), was used to conduct this study, as it has already proven its efficacy in similar contexts (Frossling et al., 2009; Hadorn et al., 2009; Knight-Jones et al., 2010; Stark, 2006; Welby et al., 2012). In the present study, key parameters such as the risk nodes affecting the probability of infection or detection were estimated using historical data as a first step. The obtained sensitivities of detection at herd and risk group level were estimated for each component. In turn, the estimated component sensitivities enabled the computation of the monthly posterior probabilities of freedom, given different probabilities of introduction. Different scenario simulations provided insight about the impact of changes in the key input parameters on the components sensitivity results.

MATERIALS AND METHODS

BTV surveillance system in Belgium in 2007

At the time of the study, 38805 cattle herds and 31777 sheep flocks were present in Belgium.

The four major components for BTV surveillance in Belgium in 2007 were:

- Passive clinical (Clinical): Clinical surveillance and notification of all ruminant case herds (sheep, cattle) (defined as positive as soon as at least one clinically infected animal was serologically confirmed in the herd);
- Sentinel (Sentinel): Monthly serological surveillance in 266 seronegative cattle herds during the vector high activity period (in 2007: April until December) to detect seroconversion (EC, 2007; FASFC, 2012);
- Winter screening (WS): Yearly cross-sectional serological survey in 344 cattle herds during the winter season (December until February) to detect if there has been past infections during the last vector season and to estimate the disease prevalence;
- Export (Export): Serological testing of all exported cattle and sheep (equivalent to 120 herds) all year round.

In each component, in case a positive serological result was found a virological isolation test was carried out.

Design of the complete disease process in a scenario tree

A scenario tree was designed for each surveillance component (WS, Sentinel, Clinical and Export) in different Excel spread sheets (Figure 1). Each risk factor affecting the probability of infection was represented by a node and a node had different branches illustrating all possible outcomes. Each outcome had a certain probability of occurrence. For each risk group, defined by the combination of nodes and branches, a detection node was entered. The risk nodes retained to influence the probability of infection at herd level were “Risk Zone” (Low/ High), “Vector Activity period” (Low/High) and “Species” (Bovine/Ovine). These were the infection risk nodes in the following study for each component. The detection node “Diagnostic Process” was characterised by the different diagnostic tests used with its corresponding sensitivity and differed for each component.

The combination of each probability determined the probability of infection and detection for each risk group. The population at risk and sampled population were partitioned according to these risk nodes and groups in order to estimate the relative effectiveness of the targeted surveillance approach of each component.

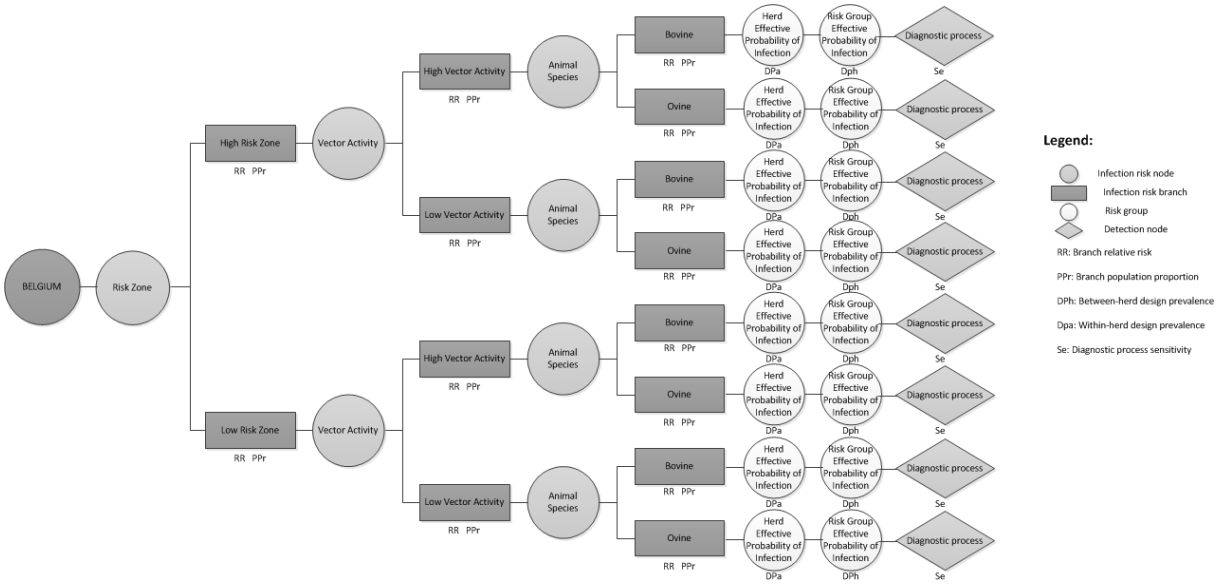


Figure 1: Scenario tree illustrating the successive events from infection to detection of BTV in Belgium

Assumptions

All surveillance system components were assumed to be independent. Information of one surveillance component was not taken into account in the other component.

The country was assumed to be free of infection. For the Sentinel, Export and Clinical components, a conservative value of 2% was used as minimum design between-herd prevalence (DPh) and within-herd prevalence (DPa) (EC, 2007). For the WS component, a 20% DPh was assumed and a 23.8 (20.1-28.1)% DPa was assumed (EC, 2007; Méroc et al., 2008). For the Clinical component, it was assumed that all animals and all herds were susceptible to be infected and show clinical signs. However, in order to take into account situations with lower disease awareness or situations where animals would no longer show clinical signs (i.e. vaccinated or naturally immunised animals), additional simulations were performed where only a fraction of the herds and animals were looked at.

The risk nodes retained to influence the probability of infection (Zone, Vector Activity period, Species) were considered independent of each other and constant over time. A specificity of 100% was assumed in each component, as each positive result was further investigated.

Table 1 Parameters characterising infection and detection nodes and branches

Node type	Node Name	Branch Description	Parameter	Data source
Infection	**Risk zone**	High risk zone: Antwerp, Brabant, Limburg, Liège	Relative Risk: Pert (1.28; 1.53; 1.98)	Collected field data: average predicted probability of infection per province
		Low risk zone: West Flanders, East Flanders, Hainaut, Namur, Luxemburg	Reference (1)	
	Vector activity period	High vector activity period: April, May, June, July, August, September, October, November, December	Relative Risk: Pert (1; 1.9; 3.14)	Collected field data: monthly proportion of infections in both periods
		Low vector activity period: January, February, March	Reference (1)	
	Animal species	Bovine	Relative Risk: Pert (1.05; 2.61; 4.16)	Literature: Durand et al., 2010
		Ovine	Reference (1)	
Detection	**Diagnostic process**	Diagnostic test Ab-ELISA (all components)	Sensitivity: Triangular (0.85; 0.89; 0.92)^{bovine} Triangular (0.78; 0.85; 0.91)^{ovine}	Literature: Vandenbussche et al., 2008
		Diagnostic test PCR (all components)	Sensitivity: Triangular (0.99; 0.99; 1)^{bovine} Triangular (0.99; 0.99; 1)^{ovine}	Literature: Vandenbussche et al., 2008
		Probability of showing clinical signs (only for passive clinical component)	Pert (0;0.025;0.116)^{bovine} Pert (0;0.077;0.097)^{ovine}	Literature: Elbers et al., 2008b
		Sensitivity of clinical signs (only for passive clinical component)	Pert (0.60;0.67;0.75)^{bovine} Pert (0.70;0.76;0.81)^{ovine}	Literature: Elbers et al., 2008c

Infection and detection nodes with their branches parameters

RISK ZONE

Spatial factors have proven to play role in the spread of BTV (Faes et al., 2012). In the present study, risk zones were identified based on historical outbreak data. Results of spatial risk factor analysis were used to define the Belgian risk zones. A logistic regression was used to model the proportion of positive farms per municipality as a function of several risk factors. The risk factors are land coverage (proportion of forest, pasture area, urban area and crop area per municipality), the altitude of a municipality, as well as the interactions between land coverage and altitude. The parameter estimates corresponding with this logistic regression model are given in appendix A.

A map, representing the predicted probabilities of infection obtained for each municipality, was produced using ArcView GIS 3.2. (ESRI). The average predicted probability of infection per municipality (1%) was used as cut-off point. The municipality was considered to be a high risk municipality if the predicted probability of infection per municipality was above 1% (red). Below the threshold of 1%, it was considered as low risk municipality (pink). Provinces borders were overlaid on the map (Figure 2). The average municipality predicted probabilities of infection per province were then computed. If it was above 1%, the province was considered as high risk zone and below as low risk zone. As a result, the provinces of Antwerp (3), Brabant (4), Limburg (5) and Liège (8) were designated as part of the high risk zone, whereas the provinces of West Flanders (1), East Flanders (2), Hainaut (6), Namur (7) and Luxemburg (9) belonged to the low risk zone. The observed minimum, maximum and most likely predicted probabilities of infection per province were used to describe the relative risk of being infected in the high risk zone branch versus low risk zone branch and fitted with a Pert distribution (Pert (1.28; 1.53; 1.98)). The low-risk zone branch, used as reference, was attributed a relative risk value of 1 (Table 1).

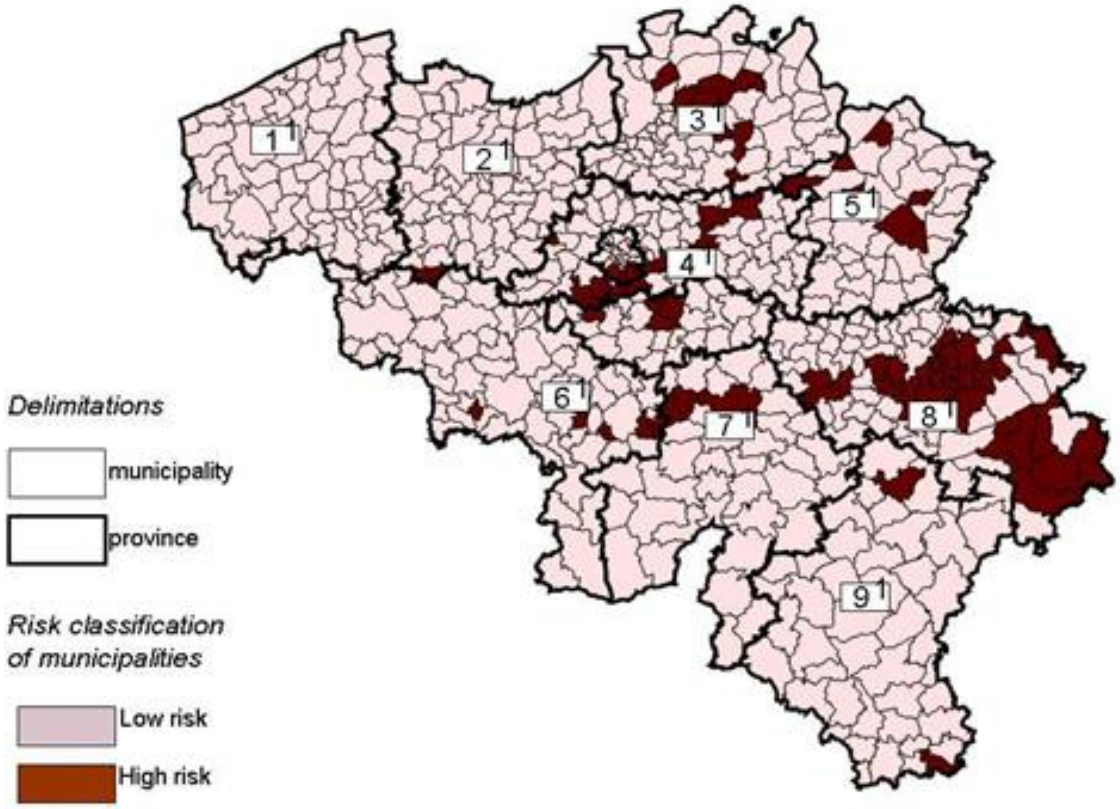


Figure 2 Map characterising the BTV risk zones in Belgium

VECTOR ACTIVITY

In 2007 in Belgium, the period from the 30th of March until the 13th of December in 2007 was considered to be the high vector activity period, whilst the remaining months of the year were considered to be in the low vector activity period (EC, 2007; FAFSC, 2012). The relative risk of being infected during the low vector activity period was used as reference, with a value of 1. To estimate the relative risk of being infected during the high vector activity period, results of export testing were explored. Export testing is carried out using serology as first line test and virological isolation as confirmation test and during the whole year. This provides some indication of the infection incidence during the high vector activity season compared to the low vector activity season. The minimum, maximum, most likely monthly proportions of positive results during the high vector activity period were used to characterise the relative risk of infection during the high vector activity period branch (Pert Distr (1; 1.9; 3.14)) (Table 1).

ANIMAL SPECIES

For two reasons cattle were considered more at risk compared to small ruminants in the present study. They are more likely to be infected (higher prevalence in cattle vs sheep) and because they exhibit less clinical signs (Elbers et al, 2008a, b; Saegerman et al., 2010), the infection could go unnoticed and spread before it is detected. The minimum, maximum and most likely values of seroprevalence proportions in both groups (bovine vs ovine) found in literature (Durand et al., 2010) were used to model the relative risk with a Pert distribution (Pert Distr (1.05; 2.61; 4.16)). The ovine branch, used as reference, was attributed a relative risk value of 1 (Table 1).

DIAGNOSTIC PROCESS SENSITIVITY

The diagnostic testing sensitivity for all components was the serial combination of both the sensitivity of the antibody ELISA test (Ab-ELISA) for BTV (ID-VET®, France) and the virological isolation PCR test (PCR) for BTV (RT-qPCR_S5), carried out as confirmation test (Vandenbussche et al., 2008). Due to the species differences, a triangular distribution (characterised by minimum, most likely and maximum values) was used for cattle for the sensitivity of the Ab-ELISA test (Triangular (0.85; 0.89; 0.92)) and for the sensitivity of the PCR (Triangular (0.99; 0.99; 1)), whereas for sheep a different triangular distribution was used for the sensitivity of the Ab-ELISA test (Triangular (0.78; 0.85; 0.91)) and for the sensitivity of the PCR (Triangular (0.99; 0.99; 1)) (Vandenbussche et al., 2008). For the Clinical component, besides the diagnostic test sensitivity of the ELISA test, the whole diagnostic process was taken into account; probability of cattle (Pert (0;0.025;0.116)) or sheep (Pert (0;0.077;0.097)) showing clinical signs (Elbers et al., 2008b); sensitivity of clinical signs for cattle (Pert (0.60;0.67;0.75)) and for sheep (Pert (0.70;0.76;0.81)) (Elbers et al., 2008c).

THE POPULATION PROPORTION AND SAMPLED PROPORTION

Table 2 represents the representative population number of herds in Belgium for each herd risk group within each component, as well as the number of sampled herds in 2007. The data were extracted from the National Animal Identification and Registration System (SANITEL) and the National Laboratory Information Management System (LIMS) of the national reference laboratory.

Table 2 Representative herd population and sampled herds in the different risk groups, defined by the combination of the infection category nodes and branches within each component (Winter screening (WS), Sentinel, Export, Clinical).

	Reference	Sampled			
Risk Group	**Population**	**WS**	**Sentinel**	**Export**	**Clinical**
HRZ/VAH/BV	14060	0	108	55	14060
HRZ/VAH/OV	11184	0	0	3	11184
HRZ/VAL/BV	14060	144	6	21	14060
HRZ/VAL/OV	11184	0	0	1	11184
LRZ/VAH/BV	24745	0	131	121	24745
LRZ/VAH/OV	20593	0	0	5	20593
LRZ/VAL/BV	24745	200	21	24	24745
LRZ/VAL/OV	20593	0	0	0	20593

HRZ/VAH/BV: High Risk Zone/Vector Activity High/Bovine LRZ/VAH/BV: Low Risk Zone /Vector Activity High/Bovine
HRZ/VAH/OV: High Risk Zone/Vector Activity High/Ovine LRZ/VAH/OV: Low Risk Zone/Vector Activity High/Ovine
HRZ/VAL/BV: High Risk Zone/Vector Activity Low/Bovine LRZ/VAL/BV: Low Risk Zone/Vector Activity Low/Bovine
HRZ/VAL/OV: High Risk Zone/Vector Activity Low/Ovine LRZ/VAL/OV: Low Risk Zone/Vector Activity Low/Ovine

Model

ESTIMATION OF SURVEILLANCE SENSITIVITIES AND POSTERIOR PROBABILITIES OF INFECTION

The relative risks (RR_{xi}) and the population proportion (PPr_{xi}) of each risk node (x) branch (i) described above enabled the estimation of the adjusted risk of infection (AR_{xi}) for each risk node branch. The obtained AR_{xi} together with the between-herd design prevalence (DPh), in turn, provided the effective probability of infection at herd level ($EPIH_j$) within each risk group (j) (defined by the combination of risk nodes (risk zone (RZ), vector activity (VA) and animal species (S)) branches) (Eq. 1 and 2).

$$AR_{xi} = \frac{RR_{xi}}{\sum_{i=1}^{2}(RR_{xi} * PPr_{xi})} \text{ (Eq. 1)}$$

$$\mathrm{EPIH_j = DPh * AR_i^{RZ} * AR_i^{VA} * AR_i^{S}} \text{ (Eq. 2)}$$

The effective probability of detection (EPD_j) for each risk group (j) within each component was derived from the different diagnostic processes, and differed according to the component considered (WS, Sentinel Export (Eq. 3) and Clinical (Eq. 4)).

$$EPD_j = Se_{DiagnosticTest_j} \text{ (Eq. 3)}$$

$$EPD_j = P_{ClinicalSigns_j} * Se_{ClinicalSigns_j} * Se_{DiagnosticTest_j} \text{ (Eq. 4)}$$

In turn these EPD_j were used to obtain the respective probabilities of detecting at least one infected animal within the tested herd (HSe_j), taking into account the average number of animals sampled (na) in each herd of average size (Na) for each risk group in each component and the within-herd design prevalence (DPa). Subsequently, the probabilities of detecting at least one infected herd within the tested risk group (GSe_j) were obtained taking into account the average number of herds sampled (nh) for risk group size (Nh) within each component in 2007, the effective probabilities of infection and the obtained herd sensitivities. Appropriate methods taking into account the sampling probability were used as described below.

If a large proportion of animals were tested within the herds, the hypergeometric approach was applied (WS, Sentinel) (Eq. 5).

$$HSe_j = 1 - (1 - (EPD_j * \frac{n_{a_j}}{N_{a_j}}))^{DPa_j * N_{a_j}} \text{ (Eq. 5)}$$

If the proportion of animals (Export) or herds (WS, Sentinel, Export) tested was smaller than 10%, the binomial approach was applied (Eq. 6,7).

$$HSe_j = 1 - (1 - (EPD_j * DPa_j))^{n_{a_j}} \text{ (Eq. 6)}$$

$$GSe_j = 1 - (1 - (HSe_j * EPIH_j))^{n_{h_j}} \text{ (Eq. 7)}$$

The exact approach was applied if all animals and herds were tested (Clinical) (Eq. 8,9).

$$HSe_j = 1 - (1 - EPD_j)^{\mathrm{D}Pa_j * N_{a_j}} \text{ (Eq. 8)}$$

$$GSe_j = 1 - (1 - HSe_j)^{EPIH_j * N_{h_j}} \text{ (Eq. 9)}$$

The estimated average number of herd and herd sizes within each component were obtained from the average values of the national animal population registers. The sampled animals and herds were obtained from surveillance results of national reference laboratory. The number of herds tested as well as the whole herd population, in each herd risk group and component over the year 2007 are shown in table 2.

The mean number of sampled animals (na) within a herd in the WS component, was set at 50 for an average herd size (Na) of 70. In the Sentinel component, na was set at 15 (in accordance with Commission Regulation (EC) No 1266/2007 (EC, 2007)). The na in the Export component was considered as 2, because on average 1 or 2 animals are exported and tested per exporting herd per year. In the Clinical

component, na was equivalent to Na or 70 as all animals were considered to be looked at.

The component sensitivities (CSe_k) provided insight on how effective each component (k) (Sentinel, Clinical, Export and WS) was for substantiating freedom of infection at a set design prevalence. The CSe_k were obtained by combining the GSe_j by the following equation (Eq. 10).

$$CSe_k = 1 - \prod_{j=1}^{8}(1 - GSe_j) \text{ (Eq. 10)}$$

In order to gain insight into the efficacy of each component in terms of early detection, monthly posterior probabilities of freedom ($PFree_{ti}$) were estimated for each component and each month (ti). The posterior probability of freedom can be obtained following Bayesian theorems. Instead of the negative predictive value of a test, the probability the country is free of infection given that the surveillance system did not detect the infection is estimated based on a prior probability of infection in the country ($PInf_{ti}$-1), the surveillance component sensitivity (CSe_{ti}) and assuming a perfect specificity (each positive results will be followed until proven to be truly negative). To begin the monthly simulations, it was assumed that no prior knowledge of the probability of infection of the previous time period (ti-1) was available; therefore a conservative value of probability of infection equal to 0.5 was considered. The latter enabled to obtain the first posterior probability of freedom after one month surveillance and thus the probability of infection (1- $PFree_{ti}$) after that first month (Eq. 11).

$$\text{PFree}_{\text{ti}} = \frac{\text{1-PInf}_{\text{ti-1}}}{\text{1-PInf}_{\text{ti-1}}*\text{CSe}_{\text{ti}}} \text{ (Eq. 11)}$$

Subsequently, the value $PFree_{ti}$ changed over time as more samples were collected. The probabilities of infection ($PInf_{ti}$) for the following months varied given the posterior probabilities of freedom of the previous months and the probability of introduction ($PIntro_{ti}$) (Eq. 12). The probability of introduction ($PIntro_{ti}$) was zero in the low vector activity period and three different values were simulated for the high vector activity period (0.25; 0.5; 0.75).

$$PInf_{ti} = (1 - PFree_{ti-1}) + PIntro_{ti} - (PIntro_{ti} * (1 - PFree_{ti-1})) \text{ (Eq. 12)}$$

The scenario trees were modelled in Microsoft Excel using @risk 5.0 software. The sensitivity estimates for the different components were obtained by separate Latin hypercube simulations for each component with 10,000 iterations in each simulation.

This provided an opportunity to consider all the possible pathways in the scenario by sampling from the parameter distributions.

SENSITIVITY ANALYSIS

To determine the input parameter that affected the component sensitivity output the most, a sensitivity analysis was carried out for each component. The obtained tornado charts allowed measuring how sensitive the output variables were to the input variables of interest, based on the correlation coefficients.

In addition different scenarios were run in each component to measure the impact of changes in the input parameters.

In all components, a scenario without relative risk was run to determine how effective the component would be without a targeted approach. For Sentinel and Export components, scenarios with higher relative risks were simulated. For the WS component, to measure how sensitive this component would be if the prevalence was lower, changes in the design prevalence (DPa and Dph of 2%) were simulated. For the Clinical component, to represent a situation with lower disease awareness, due for instance to immune animals showing less clinical signs two scenarios were simulated. The first scenario simulated only 10% of the total herd population being looked at and sampled and the second scenario simulated 50% of the total herd population.

RESULTS

Estimation of surveillance sensitivities and posterior probabilities of infection

Table 3 illustrates the respective herd and risk group sensitivities obtained for each risk group in each component after a full year of surveillance.

Table 3 Sensitivities (Mean (Minimum-Maximum)) at herd and risk group level for each risk group within each component (Winter screening (WS), Sentinel, Export, Clinical).

			WS	Sentinel	Export	Clinical
Risk Group level						
HRZ	VAH	BV		0.72(0.53-0.82)	0.09(0.06-0.11)	0.99(0.1-1)
		OV			0(0-0)	0.99(0.85-1)
	VAL	BV	1(1-1)	0.04(0.02-0.06)	0.02(0.01-0.03)	0.98(0.06-1)
		OV			0(0-0)	0.99(0.59-0.99)
LRZ	VAH	BV		0.63(0.46-0.73)	0.12(0.07-0.15)	0.99(0.11-1)
		OV			0(0-0)	0.99(0.90-1)
	VAL	BV	1(1-1)	0.08(0.05-0.12)	0.01(0.01-0.02)	0.98(0.07-1)
		OV				0.99(0.69-0.99)
Herd level						
HRZ	VAH	BV		0.25(0.24-0.26)	0.04(0.03-0.04)	0.03(0-0.09)
		OV			0.03(0.03-0.04)	0.06(0-0.09)
	VAL	BV	0.99(0.99-1)	0.25(0.24-0.26)	0.04(0.03-0.04)	0.03(0-0.09)
		OV			0.03(0.03-0.04)	0.06(0-0.09)
LRZ	VAH	BV		0.25(0.24-0.26)	0.04(0.03-0.04)	0.03(0-0.09)
		OV			0.03(0.03-0.04)	0.06(0-0.09)
	VAL	BV	0.99(0.99-1)	0.25(0.24-0.26)	0.04(0.03-0.04)	0.03(0-0.09)
		OV			0.03(0.03-0.04)	0.06(0-0.09)

HRZ/VAH/BV: High Risk Zone/Vector Activity High/Bovine LRZ/VAH/BV: Low Risk Zone /Vector Activity High/Bovine
HRZ/VAH/OV: High Risk Zone/Vector Activity High/Ovine LRZ/VAH/OV: Low Risk Zone/Vector Activity High/Ovine
HRZ/VAL/BV: High Risk Zone/Vector Activity Low/Bovine LRZ/VAL/BV: Low Risk Zone/Vector Activity Low/Bovine
HRZ/VAL/OV: High Risk Zone/Vector Activity Low/Ovine LRZ/VAL/OV: Low Risk Zone/Vector Activity Low/Ovine

The best herd and risk group sensitivities were obtained with the WS component, but only for cattle during the low vector activity period. Because no samples were taken during the high vector activity period and in sheep, no output was available for these risk groups. Sentinel component had lower herd and risk group sensitivities than WS, with the highest risk group sensitivities observed for cattle sampled during the high vector activity period. Small values were found for cattle sampled during the low vector activity period. These were actually remaining samples taken out of the vector activity period. No values were obtained for sheep risk groups as they were not sampled. The herd and risk group sensitivities were very low in the Export component, reflecting the very small number of samples taken in that component. Despite the very low herd sensitivities in the Clinical component, reflecting the very

small EPD in that component, mean risk group sensitivities were high providing that the sampled population consisted of all herds, but the range around these mean values was wide. WS, Sentinel and Clinical components appeared to be rather sensitive for detecting the infection after a whole year of surveillance with component sensitivities mean (minimum-maximum) values of 1 (1-1), 0.9 (0.77-0.95) and 1 (0.99-1) respectively, whilst Export only had very low sensitivities of detection 0.22 (0.17-0.26). In terms of early detection (Figure 3), the Clinical component appeared to be the most efficient for guaranteeing posterior probability of infection freedom. The WS component was also very efficient, and despite the fact that samples were only taken during the month of January, the posterior probability remained high until the probability of introduction changed which then caused a decreasing posterior probability of freedom. The posterior probability of freedom with the Sentinel component was equal to the prior probability of infection 0.5 during the low vector activity period as no samples were taken. It then suddenly increased in March, but only shortly. As probability of introduction increased the posterior probability of freedom decreased. The Export component only provided very limited guarantee towards the country posterior probability of freedom and only during the low vector activity period, thereafter the posterior probability of freedom decreased drastically.

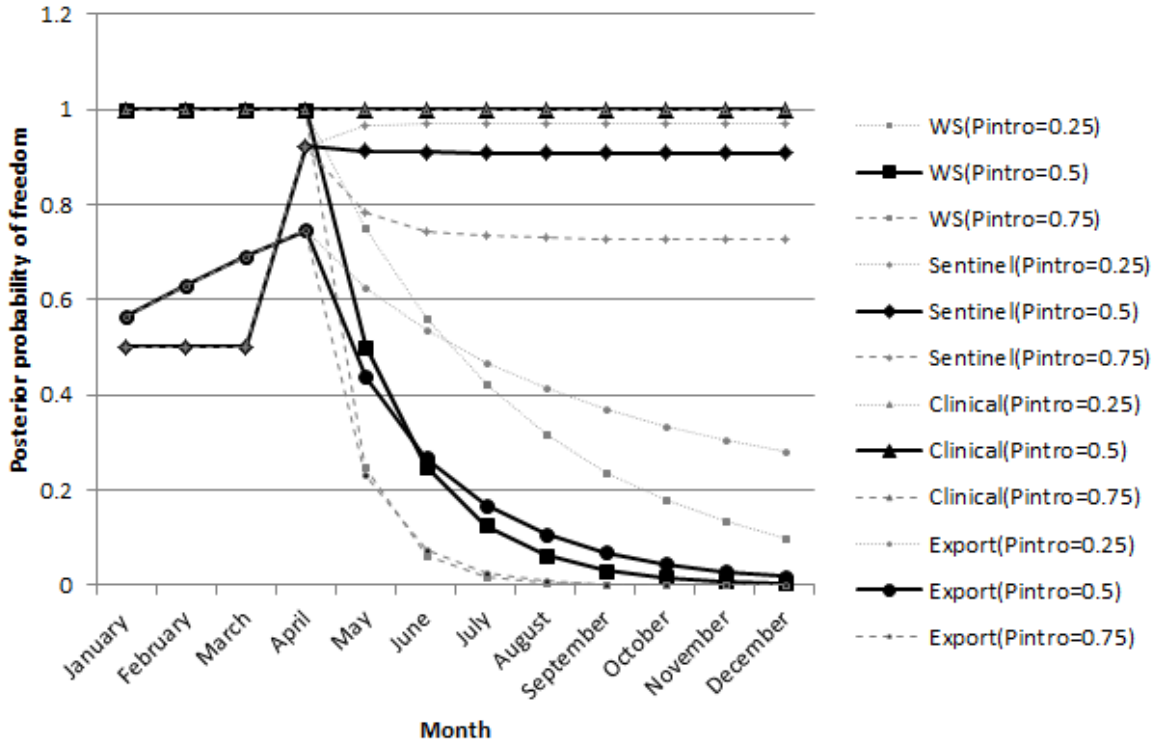
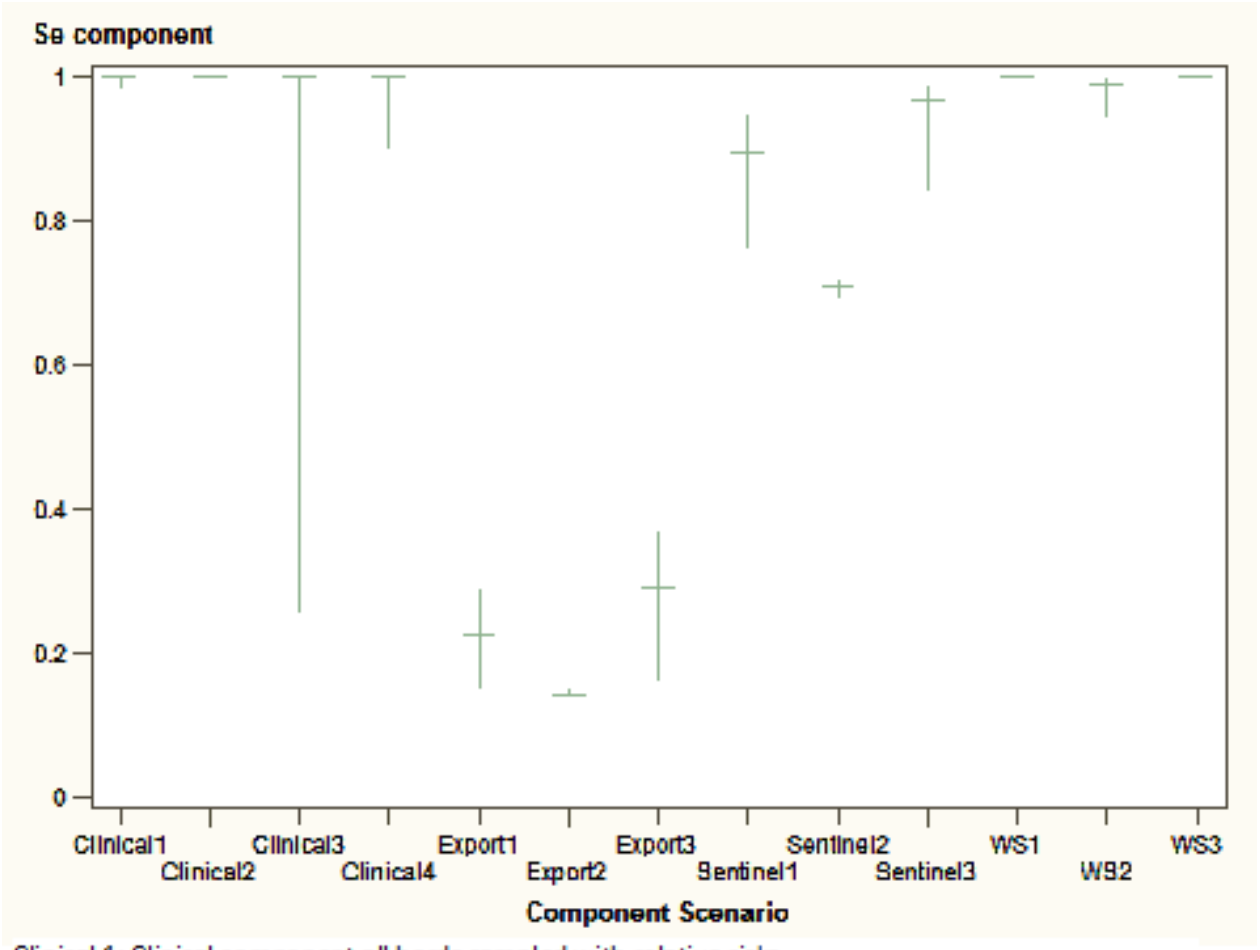


Figure 3 Monthly posterior probabilities of disease freedom for each component

Sensitivity analysis

For the Clinical component the diagnostic process was the parameter that influenced mostly the component sensitivity. For WS, the most influential parameters were the within-herd design prevalence followed by the diagnostic test sensitivity and the relative risks input values for the infection nodes vector activity and species. For the Export and Sentinel component, the sensitivity analysis showed that the most influential inputs were the relative risks values attributed to vector activity and species infection nodes followed by the diagnostic test sensitivities. The range of the rank correlations values differed in each component. The highest impact was for the WS (rho= 0.9) followed by Sentinel and Export component (rho= 0.7) and Clinical (rho=0.6),

The scenario simulations enabled to further estimate the impact of the input values (Figure 4). For each component, the sensitivity was estimated through a simulation ignoring differential risk of infection in the different risk nodes. This resulted in a drop of 10-20% component's sensitivities for Export and Sentinel component only. If higher relative risks values were considered sensitivities raised in those same components, reflecting the targeted approach of these components. For WS, if it was assumed that the disease was only present at 2% between-herd and within herd prevalence, the sensitivity of this component dropped from 1 to 0.99. In the Clinical component the estimated sampled population had a big impact. Indeed if only 10% of the herds were sampled the sensitivity varied much more, with a minimum value of only 0.26. The decrease in sensitivity was less evident if 50% were sampled with minimum values of 0.9.



Clinical 1: Clinical component all herds sampled with relative risks
Clinical 2: Clinical component all herds sampled without relative risks
Clinical 3: Clinical component with 10% of herds sampled with relative risks
Clinical 4: Clinical component with 50% herds sampled with relative risks
Export 1: Export component all herds sampled with relative risks
Export 2: Export component all herds sampled without relative risks
Export 3: Export component all herds sampled with higher relative risks
Sentinel 1: Sentinel component all herds sampled with relative risks
Sentinel 2: Sentinel component all herds sampled without relative risks
Sentinel 3: Sentinel component all herds sampled with higher relative risks
WS 1: WS component all herds sampled with a relative risks, between and within herd design prevalence (0.2; 0.24) respectively
WS 2: WS component all herds sampled with relative risks, between and within herd design prevalence (0.02; 0.02) respectively
WS 3: WS component all herds sampled without relative risks, between and within herd design prevalence (0.2; 0.24) respectively

Figure 4 Sensitivity analysis: Impact of changes in input parameters on the component sensitivity

DISCUSSION

The aim of the present study was to estimate the effectiveness of the Belgian surveillance system for bluetongue and its components in terms of substantiating freedom of infection, early detection and their risk-based approach. Results provided a good insight into the effectiveness of the targeted approach of each component. The scenario tree methodology, used in the present study, offers the opportunity to take into account all factors that can influence the probability of infection and detection while estimating the sensitivity of detection of a surveillance program, providing that information about key input parameters influencing these probabilities is available. Assumptions are used where information about those key input parameters is missing. This might be a source of bias (Dohoo et al., 2009). In the present study, assumptions regarding input parameters were mitigated by using field data from the previous year's outbreaks. Risk groups where infection was most likely to cluster, were identified. Subsequently, this study provided a clear indication of the different herd and risk group sensitivities in the different surveillance components. The sensitivity analysis allowed measuring the impact of changes in the input parameters which provided further insight on the efficacy of each component in terms of early detection and freedom of infection substantiation.

Overall high sensitivities at herd level were obtained for the WS and Sentinel components whilst this was not the case for the Clinical and Export component. The low diagnostic process sensitivity in the Clinical component explains the low sensitivities at herd level. In the Export component because only few animals within a herd were sampled, despite the good diagnostic test sensitivities, sensitivities at herd level remain low. When looking at the sensitivities at risk group level, the Clinical and WS component had higher mean sensitivities values compared to all other component. However, the Clinical component had a very wide range around the mean risk group sensitivity values. In terms of early detection and guaranteeing the posterior probability of freedom of infection level, the Clinical component turned out to be most effective, followed by the Sentinel component. Despite the very low herd sensitivity values in the Clinical component (Table 3), the large number of sheep and cattle herds processed monthly during the whole year enabled maintaining a high level of posterior probability of freedom throughout the year. Simulations of different

scenarios were carried out to measure the impact of the assumptions used in the model regarding sampled population, relative risks and the design prevalence.

In a situation where disease awareness would drop down or if clinical signs would no longer be visible, as in a immunised population or with a different serotype then BTV-8, the assumption that all herds and animals are looked at and sampled might no longer hold.

Two different scenarios were run (Figure 4) to estimate the impact of the sampled population within the passive clinical component. The first one, where only 10% of the herds were looked at and sampled, made the sensitivity at component level drop down significantly with minimum values up to 0.26. When only 50 % of the herds were looked at and sampled, the sensitivity was acceptable but not as good as when 100% of the herds were considered. Not only is passive clinical surveillance strongly dependent on disease awareness and the occurrence of clinical symptoms, but also farmers could be reluctant to report diseases by fear of ethical and economic repercussions (Elbers et al., 2010; Saegerman et al., 2010). More in depth study of this parameter is required in order to better estimate the sensitivity of this component. Furthermore, in a situation where vaccination is applied or where natural immune status has been established following past recurrent infections, clinical signs may no longer appear which may result in a decrease in disease awareness. In these situations case, a laboratory testing program or a syndromic surveillance probably would be more appropriate.

If no relative risks were considered, component sensitivities dropped down for Sentinel and Export component only.

The WS component, which displayed very good sensitivity at cattle herd and risk group level during the whole low vector activity period, was evaluated with a between-herd design prevalence of 20% and average of 23.8% within-herd design prevalence, as requested by the EU legislation. However, to determine the effectiveness of this component in terms of substantiating disease freedom, it was thought that 2% design prevalence would be interesting to simulate at within- and between-herd level. Results showed that though sensitivities at risk group level would be lower, the component sensitivity would be 0.99 instead of 1. This would only result in a slight decrease. Because this component consists only of yearly cross sectional survey, it is of limited value for early detection. Thus it may be concluded that WS is

useful for substantiating freedom of infection at the end of a year or for sero prevalence estimation.

The Sentinel component showed relatively good sensitivities at herd and risk group level, despite the fact that not all herds and animals were sampled within this component. However, in terms of early detection, due to the limited number of herds followed-up and the high probability of introduction in the high vector activity period, the Sentinel component might not be optimal for detecting a 2% seroconversion and substantiating freedom of infection.

Export testing only had limited value in Belgium due to the small number of samples taken in this component. Thus relying only on export testing, in Belgium, is not sufficient to substantiate infection freedom. However, it should be stated that this was not the primary aim of export testing.

CONCLUSIONS

Some recommendations for the future BTV surveillance in Belgium can be made from this study's output;

- A Sentinel program is very effective to detect the infection, provided that sufficient samples are taken and that the sampling frequency is high enough. A monthly or quarterly frequency would be recommended.
- WS is useful for substantiation of freedom an overall prevalence interpretation at the end of the year.
- Export testing by itself is not enough to guarantee freedom of infection or early detection of infection.
- Clinical passive surveillance is effective for both freedom of infection substantiation or early detection of infection but dependent on a few constraints.

This study has enabled to better quantify the sensitivity of the main surveillance components and their risk-based approach for BTV detection in Belgium. Simulations carried out provided insight into the impact of several assumptions, which further enabled better quantification of the effectiveness of the different bluetongue surveillance components as required by the European legislation and the international standards.

Appendix A

Logistic regression parameter estimates of the different risk factors used to model the proportion of positive farms per municipality

Effect	*Estimate*	*Error*	*DF*	*T Value*	*Pr > |t|*
Intercept	-7.3913	0.9203	9947	-8.03	<.0001
P_FOREST	0.1740	1.1775	9947	0.15	0.8825
P_PASTURE	-0.9095	1.1217	9947	-0.81	0.4175
P_URBAN	2.1159	1.6069	9947	1.32	0.1879
P_CROP	1.7920	0.9416	9947	1.90	0.0570
altitudemean	0.0065	0.0012	9947	5.38	<.0001
P_FOREST*P_PASTURE	15.8376	2.1863	9947	7.24	<.0001
P_FOREST*P_URBAN	14.8175	3.0954	9947	4.79	<.0001
P_PASTURE*P_URBAN	12.6241	3.0000	9947	4.21	<.0001
P_URBAN*P_CROP	6.4707	2.0652	9947	3.13	0.0017
altitudemea*P_FOREST	-0.0121	0.0025	9947	-4.84	<.0001
altitudemean*P_CROP	-0.0071	0.0019	9947	-3.67	0.0002
Residual	2.0009	.	.	.	.

REFERENCES

Dohoo I, Martin W, Stryhn H (Eds.), 2009. Targeted sampling. In: Veterinary Epidemiologic Research. VER Inc, Canada., 41-42.

Durand B, Zanella G, Biteau-Coroller F, Locatelli C, Baurier F, Simon C, Le Dréan E, Delaval J, Prengère E, Beauté V, Guis H, 2010. Anatomy of bluetongue virus serotype 8 epizootic wave, France, 2007-2008. Emerg Infect Dis. 16(12): 1861-1868.

European Commission (EC), 2007. Commission Regulation 1266/2007/EC of 26 October 2007 on implementing rules for Council Directive 2000/75/EC as regards the control, monitoring, surveillance and restrictions on movements of certain animals of susceptible species in relation to bluetongue. Official Journal of the European Union L283, 37-52.

Elbers ARW, Backx A, Meroc E, Gerbier G, Staubach C, Hendrickx G, van der Spek A, Mintiens K, 2008a. Field observations during the bluetongue serotype 8 epidemic in 2006. I. Detection of first outbreaks and clinical signs in sheep and cattle in Belgium, France and the Netherlands. Prev Vet Med. 87(1-2): 21-30.

Elbers ARW, Backx A, Mintiens K, Gerbier G, Staubach C, Hendrickx G, van der Spek A, 2008b. Field observations during the Bluetongue serotype 8 epidemic in 2006. II. Morbidity and mortality rate, case fatality and clinical recovery in sheep and cattle in the Netherlands. Prev Vet Med. 87(1-2): 31-40.

Elbers ARW, Backx A, Ekker HM, Van der Spek AN, Van Rijn PA, 2008c. Performance of clinical signs to detect bluetongue virus serotype 8 outbreaks in cattle and sheep during the 2006-epidemic in The Netherlands. Vet Microbiol. 129: 156-162.

Elbers ARW, Gorgievski-Duijvesteijn MJ, Zarafshani K, Koch G, 2010. To report or not to report: a psychosocial investigation aimed at improving early detection of avian influenza outbreaks. Rev sci tech Off Int Epiz. 29(3): 435-449.

Faes C, van der Stede Y, Guis H, Staubach C, Ducheyne E, Hendrickx G, Mintiens K, 2012. Factors affecting Bluetongue serotype 8 spread in Northern Europe in 2006: The geographical epidemiology. Prev Vet Med. pii: S0167-5877(12)00403-5. doi: 10.1016/j.prevetmed.2012.11.026. [Epub ahead of print]

Federal Agency for the security of the food chain (FASFC): Fièvre catarrhale du mouton, Mesures et procédures, document E 605. URL (Last update date: 19/12/2012): http://www.favv.be/santeanimale/fievrecatarrhale/mesures.asp (Consulted on: 13/05/2013).

Frossling J, Agren EC, Eliasson-Selling L, Lewerin SS, 2009. Probability of freedom from disease after the first detection and eradication of PRRS in Sweden: scenario-tree modelling of the surveillance system. Prev Vet Med. 91: 137-145.

Hadorn DC, Racloz V, Schwermer H, Stark KDC, 2009. Establishing a cost-effective national surveillance system for Bluetongue using scenario tree modelling. Vet Res. 40(6): 57.

Knight-Jones TJ, Hauser R, Matthes D, Stark KD, 2010. Evaluation of effectiveness and efficiency of wild bird surveillance for avian influenza. Vet Res. 4, 41(4): 50.

Martin PAJ, Cameron AR, Greiner M, 2007a. Demonstrating freedom from disease using multiple complex data sources 1: a new methodology based on scenario trees. Prev Vet Med. 79: 71-97.

Martin PAJ, Cameron AR, Barfod K, Sergeant ESG, Greiner M, 2007b. Demonstrating freedom from disease using multiple complex data sources 2: case study--classical swine fever in Denmark. Prev Vet Med. 79: 98-115.

Mellor PS, Carpenter S, Harrup L, Baylis M, Mertens PP, 2008. Bluetongue in Europe and the Mediterranean Basin: history of occurrence prior to 2006. Prev Vet Med. 87: 4-20.

Méroc E, Faes C, Herr C, Staubach C, Verheyden B, Vanbinst T, Vandenbussche F, Hooyberghs J, Aerts M, De Clercq K, Mintiens K, 2008. Establishing the spread of Bluetongue virus at the end of the 2006 epidemic in Belgium. Vet Mic. 131: 133-144.

Saegerman C, Mellor P, Uyttenhoef A, Hanon JB, Kirschvink N, Haubruge E, Delcroix P, Houtain JY, Pourquier P, Vandenbussche F, Verheyden B, De Clercq K, Czaplicki G, 2010. The most likely time and place of introduction of BTV8 into Belgian ruminants. PLoS One 5(2): e9405.

Stark KDC, Regula G, Hernandez J, Knopf L, Fuchs K, Morris RS, Davies P, 2006. Concepts for risk-based surveillance in the field of veterinary medicine and veterinary public health: review of current approaches. BMC Health Serv Res. 6: 20.

Vandenbussche F, Vanbinst T, Verheyden B, Van Dessel W, Demeestere L, Houdart P, Bertels G, Praet N, Berkvens D, Mintiens K, Goris N, De Clercq K, 2008. Evaluation of antibody-ELISA and real-time RT-PCR for the diagnosis and profiling of bluetongue virus serotype 8 during the epidemic in Belgium in 2006. Vet Microbiol. 129: 15-27.

Welby S, Govaerts M, Vanholme L, Hooyberghs J, Mennens K, Maes L, Van Der Stede Y, 2012. Bovine tuberculosis surveillance alternatives in Belgium. Prev Vet Med. 106(2): 152-161.

Wilson AJ, Mellor PS, 2009. Bluetongue in Europe: past, present and future. Phil. Trans. R. Soc Lond B Biol Sci. 364: 2669-2681.

CHAPTER VII: EFFECTIVENESS AND COST EFFICIENCY OF PASSIVE AND ACTIVE SURVEILLANCE COMPONENTS FOR EARLY DETECTION AND PROOF OF FREEDOM OF EMERGING VECTOR-BORNE DISEASES IN CATTLE: BLUETONGUE AS CASE STUDY FOR BELGIUM, FRANCE AND THE NETHERLANDS

Transbound Emerg Dis., 2016. doi: 10.1111/tbed.12564. [Epub ahead of print]

Cost efficiency of passive active and syndromic surveillance components for early detection of emerging vector borne disease in cattle: Bluetongue as case study for Belgium, France and the Netherlands.

Welby S, van Schaik G, Veldhuis A, Brouwer H, Peroz C, Berends I, Fourichon C, Wever P, van der Stede Y.

Abstract

Effective and efficient early detection of emerging vector-borne disease as well as quick recovery of country's freedom status remain a constant challenge in animal health surveillance. In the present study, the original scenario tree methods developed by Martin et al. in 2007 were adapted in order to allow comparison of the relative efficacy and cost efficiency of different surveillance components in proving the absence of infection (freedom of infection) or early detection of vector-borne disease in cattle populations within different countries. Bluetongue outbreak and surveillance data from the epidemic that occurred in the Netherlands, France and Belgium, in the years 2006 and 2007 provided a reliable input data for evaluating the different surveillance components in place in each country. In a first stage, the different surveillance components (sentinel, yearly cross sectional and passive clinical reporting) in place in 2006-2007 within each country were evaluated in terms of efficacy for substantiating freedom of infection. The yearly cross sectional survey and passive clinical reporting within each country performed well with sensitivity of detection values ranging around 0.99. The sentinel component had a sensitivity of detection around 0.7. The large number of cattle and herds sampled and the high diagnostic tests sensitivity contributed to these high sensitivity values within each of these surveillance components. In a second stage, it was investigated how effective the components were for (early)-detection and if syndromic surveillance using reproductive performance data, milk production and mortality data could be of added value using production data from the Netherlands and Belgium that were available from 2006-2007. To account for the timeliness of detection, epidemic curves obtained from Reed-Frost models were used to estimate the input to feed the scenario tree models. Depending on the expected within herd prevalence, passive clinical reporting and syndromic surveillance performed much better than the active components, with respectively 0.99-1.

Depending on the assumed direct costs of running each of these surveillance components, passive clinical surveillance together with syndromic surveillance (based on reproductive performance data) turned out most cost efficient for detection of vector-borne diseases such as Bluetongue.

To conclude, for emerging or re-emerging vector-borne disease such as Bluetongue, it is recommended to use passive clinical and syndromic surveillance as early detection systems for maximum cost efficiency. Once an infection is detected, a cross sectional screening is recommended for substantiating freedom of infection. Sentinel surveillance is useful for monitoring the disease evolution providing sufficient herds are sampled.

INTRODUCTION

In a context of globalisation and climate change, emerging and re-emerging diseases constitute a constant threat to animal health. During the years 2006 and 2007, an arthropod-borne viral disease, named Bluetongue, emerged in Northern Europe with major economic impacts and adverse impact on trade (Wilson and Mellor 2009; Zientara and Sánchez-Vizcaíno, 2013). Early detection and reestablishment of freedom of infection were important goals to ensure rapid control of disease as well as national and international trade. National and international legislation standards set the minimum legal requirements for early detection and recovery of freedom status but it is also up to each country to prove the effectiveness of their surveillance strategies (FAO, 1999; OIE, 2014). In this context and following the major economic consequences of the Bluetongue outbreaks, regulation 1266/2007/EC last amended by 456/2012/EC (EC, 2012) prescribes the main surveillance components to implement in the cattle population within each country. These are active surveillance components (mainly characterized by yearly cross sectional surveys or sentinel serological/virological screening to detect seroconversion) and passive surveillance components such as clinical reporting. The minimum guidelines for sample size are laid down in the regulation together with the maximum design prevalence as well as desired confidence level.

Martin et al., (2007) proposed interesting tools to evaluate surveillance systems that have been used extensively in numerous studies for different diseases (endemic or epidemic) (Frössling et al., 2013; Knight-Jones et al., 2010; Welby et al., 2013). These methods offer the advantage of combining several data sources when evaluating a national surveillance program (ie slaughterhouse surveillance, purchase

testing, bulk milk data, etc…). It also enables to take into account, not only the sensitivity of the diagnostic test used, but also the whole diagnostic process, together with the number of herds, cattle present and tested in the population, the expected prevalence and differential risk of infection (if present). This provides a very powerful and objective tool for evaluating surveillance and obtaining reliable values about the effectiveness of a surveillance system in terms of guaranteeing freedom of infection. However, it requires, as many other statistical or mathematical models, the availability of data. When the input data are missing, use of literature values or assumptions are a possibility, but care must be taken to avoid bias (Dohoo et al., 2009). Another limitation of the present tool is that it does not enable the evaluation in terms of early detection. In the present study, we propose an alternative approach to account for timeliness of detection using data obtained with epidemic curves from Reed-Frost models as inputs to feed the scenario models.

The aim of this work was three fold:

- First, the different surveillance components in place in Belgium, France and the Netherlands for Bluetongue during 2006 and 2007 were evaluated in terms of effectiveness for substantiating freedom of infection using the scenario tree methods;
- In a second stage, effectiveness in terms of early detection of the existing surveillance components and the added value of using syndromic surveillance (based on mortality, milk production or reproductive performance data) was estimated combining methods of Reed-Frost and scenario tree models.
- Finally, cost efficiency of each surveillance component was investigated.

MATERIAL AND METHODS

Surveillance components

Bluetongue surveillance data at the start of the epidemic in 2006 and 2007 were used for this study. The surveillance components conducted within each country during 2006 and 2007,as well as syndromic surveillance, not prescribed by the European regulation 1266/2007, and considered in the present study are defined in Table 1.

Table 1: Definition of each surveillance component considered in the present study.

Component	***Description***	***Regulatory framework***
Cross Sectional	Yearly cross sectional survey of a sample from the total cattle population during the winter	Regulation 1266/2007/EC (last amended by 456/2012/EC
Sentinel	Monthly follow up of samples of seronegative cattle in order to detect seroconversion in the months were vectors are active	Regulation 1266/2007/EC (last amended by 456/2012/EC
Passive clinical	Immediate reporting of relevant clinical signs by the farmers and vets to the authorities	Regulation 1266/2007/EC (last amended by 456/2012/EC
Syndromic surveillance: - milk production - reproductive performance - mortality	Defined by the collection and analysis of nonspecific production data (ie milk production data; reproductive performance data; mortality data) to detect deviations of the baseline trends	None

Population data and samples taken within the different surveillance components for BTV in 2006-2007 (Cross Sectional, Sentinel, Passive Clinical, Milk production data, Reproductive performance data, Mortality data) are summarised for each country ((BE (Belgium); FR (France); NL (the Netherlands)) in Table 2. Population and surveillance data from official programs were obtained from the different regional and national laboratories in the given countries (Veterinary and Agriculture Research Center (CODA) in Belgium; Oniris Unité de recherche Biologie, Epidémiologie et analyse de risque en Santé Animale – BioEpAR in France; GD Animal Health in the Netherlands) and supported by literature (Méroc et al., 2008; Satman-Berends et al., 2010; van Schaik et al., 2008). Syndromic surveillance data based on monthly milk production data and reproductive performance data were obtained from the Royal Cattle Syndicate (CRV). CRV handles monthly milk recording data of 80-85% of Dutch and approximately 48% of Flemish dairy herds in Belgium. An extrapolation to the whole country of Belgium was made based on the cattle population data.

Mortality data were obtained from rendering plants (RENDAC) in Belgium and the Netherlands.

Table 2: Cattle population and surveillance data for each surveillance component in place in each country ((BE (Belgium); FR (France); NL (the Netherlands)) for Bluetongue in 2006-2007 (NA: not available).

Country	Data Level	Population	Cross-sectional	Sentinel	Passsive clinical	Milk production data	Reproductive performance data	Mortality data
BE	**Herd**	36000	344	260	ALL	10000	10000	7500
	Within herd	50(1-370)	70%	15	ALL	70%	11(7-14)	2(1-2)
FR	**Herd**	190000	9300	960	ALL	NA	NA	NA
	Within herd	100	10(1-50)	15	ALL	NA	NA	NA
NL	**Herd**	44204	5436	275	ALL	15500	19270	9923
	Within herd	78(1-383)	1(1-12)	16	ALL	70%	11(7-14)	2(1-2)

To account for the diversity in herd sizes, a distribution was used to model the within herd population with mean (25th percentile-75th percentile) of herd sizes. For cross sectional survey in Belgium and milk production data in both countries, 70% of the herd size was considered as it was thought that only 70% would be lactating cows and thus considered for sampling. For reproductive performance data and mortality data, the varying within herd sampled population mean (min-max) values was accounted for. All cattle within herd were considered to be looked at in the passive clinical component and a fixed number of 15 or 16 cattle, as stated by the regulation, were considered for the sentinel component.

Effectiveness of surveillance in terms of freedom of infection

Scenario trees, modelling the sensitivity of detection at herd and component level, were designed for each bluetongue ongoing surveillance component (sentinel, yearly cross sectional and passive clinical reporting) in 2006-2007 within each country (France, the Netherlands and Belgium) in Microsoft Excel spreadsheets. Effectiveness in terms of substantiating infection freedom was estimated based on the probability of detecting an infected animal or herd, in other words sensitivity, given a set design prevalence. Herd (HSe) (Eq.1) and component sensitivity (CSe) (Eq.2), were computed for each surveillance component (i) for substantiating freedom of infection in each country, using identical hypergeometric sampling

approaches. Simulations were done using ModelRisk® software with 10.000 iterations per simulation. Appropriate distributions were fitted to the data to account for the uncertainty and variability of the different parameters.

$$HSe_i = 1 - \left(1 - \left(EPD_i * \frac{n_{a_i}}{N_{a_i}}\right)\right)^{N_{a_i} * DP_{a_i}} \text{(Eq. 1)}$$

$$CSe_i = 1 - \left(1 - \left(HSe_i * \frac{n_h}{N_{h_i}}\right)\right)^{N_{h_i} * DP_{h_i}} \text{(Eq. 2)}$$

Average herd (Nh) and average within herd (Na) population data as well as surveillance data regarding average number of herds tested (nh) and average animals tested per herd (na) were obtained from national cattle identification databases and regional/national laboratories information systems within each country (Table 2). The expected design prevalence at herd (DPh) and within herd (DPa) level was set at 2% (according to EU regulation 1266/2007/EC). The effective probabilities of detection (EPD) for the active surveillance components (Cross sectional and Sentinel) were minimum, most likely, maximum sensitivity values of the antibody ELISA test (Ab-ELISA) for BTV (ID-VET®, France) (Vandenbussche et al., 2008) and characterised by a pert distribution (Pert (0.85; 0.89; 0.92)). Because no data was available for the EPD sensitivities for passive clinical surveillance component, the diagnostic process minimum, most likely, maximum sensitivity values was modeled with a pert distribution (Pert (0.1; 0.5; 0.99)).

Effectiveness of surveillance in terms of early detection

In terms of early detection, the model described by Martin et al. (2007), was adapted to account for timeliness of detection. Also the added value of syndromic surveillance using routinely collected production data (milk production data, reproductive performance data (gestational length-indicators), and mortality data (obtained from rendering plants)) was evaluated in case it were to be in place. The EPD values were adjusted using the number of detected cases over the number of expected infected cases obtained with epidemic curves. The epidemic curves were obtained using between herd reproductive ratios for BTV in 2006 and 2007 obtained from literature (de Koeijer et al., 2011; Santman-Berends et al., 2013) and simulated by Reed Frost Models (RFM) in Winepiscope 2.0.®, with monthly time intervals. Different epidemic curves were obtained using different between herd reproductive ratios (R-values)

(R=1; R=2; R=3; R=4 and R=5). The number of cases, infected and susceptible cattle were obtained for each month from April 2006 (time at which the bluetongue virus was assumed to be introduced in Belgium in an area near the Dutch border (Saegerman et al., 2011) until December 2006 (time at which the first epidemic wave stopped). Subsequently, the same exercise was done for 2007 (April to December). For that period the susceptible and infected population, remaining at the end of 2006, was used as starting point to simulate the RFM. The number of detected infected cattle at time of first alerts observed with milk production data, reproductive performance data, mortality data or dates of first case notification with the sentinel, cross sectional and passive clinical component separately were used to compute the EPD (Eq. 3).

$$EPD_i = \frac{1(\text{detected})_{\,month\ of\ first\ alert_i}}{\text{N (infected)}_{\,month\ of\ first\ alert_i}} \qquad \text{(Eq. 3)}$$

The number of detected cattle was always one and corresponded to the first case identified or notified at time of first alert. The number of infected cattle in the same month was obtained from the RFM output. This was conducted for each simulated epidemic curve obtained from the different transmission ratios (R-values). Finally the mean, minimum and maximum of the different EPD values obtained for each simulated R-value were used for characterising the distribution of the EPD value of each component. The same exercise was done for both countries. The dates of first alerts triggered by syndromic surveillance were obtained from a study of Veldhuis et al., (submitted) and the dates of first cases by Passive Clinical, Sentinel and Cross Sectional surveillance were obtained using historic laboratory data (CODA and GD). For example, the passive clinical surveillance component in Belgium detected the first case in August 2006, therefore to obtain the EPD of the passive clinical component, we used 1 over the expected infected population in the month of August 2006 simulated by the RFM (Méroc et al., 2008; Santman-Berends et al., 2010). Similarly, the first alerts with syndomic surveillance based on reproductive performance data in Belgium were estimated to be observed in August 2007, therefore we used 1 over the expected infected population in the month of August 2007 to estimate the EPD of that syndromic surveillance component.

The effective probabilities of detection (EPD) mean (min-max) values obtained for each epidemic curve simulated are summarised in Table 3. The herd and component sensitivities were obtained using eq. 1 and 2. Because the legislation for BTV only prescribes design prevalence at animal level, two different simulations were considered to estimate the impact of considering a clustering effect at herd level. In each simulation the herd prevalence was held constant at a value of 2%, whereas the within herd prevalence varied between 2 (simulation 1) and 20% (simulation 2) in order to measure the impact of the expected within herd design prevalence on the effectiveness of detection.

Table 3: Parameters regarding effective probabilities of detection (EPD) mean (min-max) values, for Belgium (BE) and the Netherlands (NL) in each surveillance component.

Component	***BE***	***NL***
EPD Cross Sectional	(0.00003;0.02266;0.11111)	(0.00002;0.02265;0.11111)
EPD Sentinel	(0.00003;0.0171;0.08333)	(0.00002;0.01709;0.08333)
EPD Passive Clinical	(0.00641;0.07197;0.25)	(0.01176;0.08836;0.25)
EPD Milk production data	(0;0;0)	(0.00027;0.03726;0.16667)
EPD Reproductive performance data	(0.00003;0.01581;0.07692)	(0.00002;0.0158;0.07692)
EPD Mortality data	(0;0;0)	(0.00007;0.03038;0.14286)

No significant deviation from the baseline trend were identified using reproductive performance data and mortality data in Belgium

Posterior probabilities of Freedom

The monthly posterior probabilities of freedom (PFree) were obtained for each component (i) (Eq. 4, 5) taking into account the probability of introduction (Pintro), probability of infection (PInf) at each time period (t=month) and the component sensitivity (CSe):

$$PFree_{i_t} = \frac{1-PInf_{i_{t-1}}}{1-PInf_{i_{t-1}} * CSe_i} \text{(Eq. 4)}$$

$$PInf_{i_t} = \left(1 - PFree_{i_{t-1}}\right) + PIntro_{i_t} - \left(PIntro_{ii_t} * \left(1 - PFree_{i_{t-1}}\right)\right) \text{ (Eq. 5)}$$

As no prior knowledge regarding the probability of introduction per month was available we assumed a conservative value of 0.5 as probability of introduction and constant through the year. Only the sensitivity values for those months where the

components were effective were considered (i.e. yearly cross sectional only carried out during the winter). During the remaining months the component sensitivity was considered null.

Cost efficiency

In a last stage, the cost of each component for early detection in Belgium was estimated, which enabled quantification of the economic efficiency of each component (Table 4) including syndromic surveillance if it were to be in place. Because cost data was only available for Belgium, the exercise regarding cost efficiency was only conducted for Belgium.

Table 4: Costs estimation of the surveillance activities needed in the different components

Components	***Surveillance activity***	***Source of information***	***Cost***	***Common to all activities***
Cross sectional **Sentinel** **Passive Clinical**	**ELISA cost**	Belgian Federal Agency for the security of the food chain	4€/test	**Development and Overall management** **Hard and software**
	Sample preparation		0.6€/test	
	Sample collection		2.5€/test	
	Farm visit by vet		25€/visit	
	Data analysis	Belgian National Laboratory: 1 person month	40000€/year	
Milk production data **Reproductive performance data** **Mortality data**	**Data deliveries**	Belgian National Laboratory: Data cost for 5 retrospective years	8000€/ year	
	Data analysis	Belgian National Laboratory: 3 person month	120000€/year	

For each ongoing active and passive surveillance component (sentinel, cross sectional, passive clinical), the cost estimation over the year consisted of a farm visit by the vet (times number of farms visited yearly) and sample collection, preparation and analysis (times number of animals sampled in each visited farm yearly). For the

syndromic surveillance components (milk production data, reproductive performance data, mortality), the considered costs were estimated as follows: the cost of data deliveries, and the cost of data analysis and interpretation of signals. The cost of hardware and software development as well as overall management was assumed common and constant to all surveillance systems and thus not accounted in the cost efficiency calculations. The yearly cost efficiency ratio was obtained using the mean component sensitivity over the total estimated cost.

RESULTS

Effectiveness of surveillance in terms of freedom of infection

Figures 1 and 2 illustrate the iteration results (1st-50th-99th percentiles) for the obtained sensitivity values (probability of detecting an infected cow or herd given it is infected at the set design prevalence) obtained at herd and component level for each the surveillance components in place in 2006-2007.

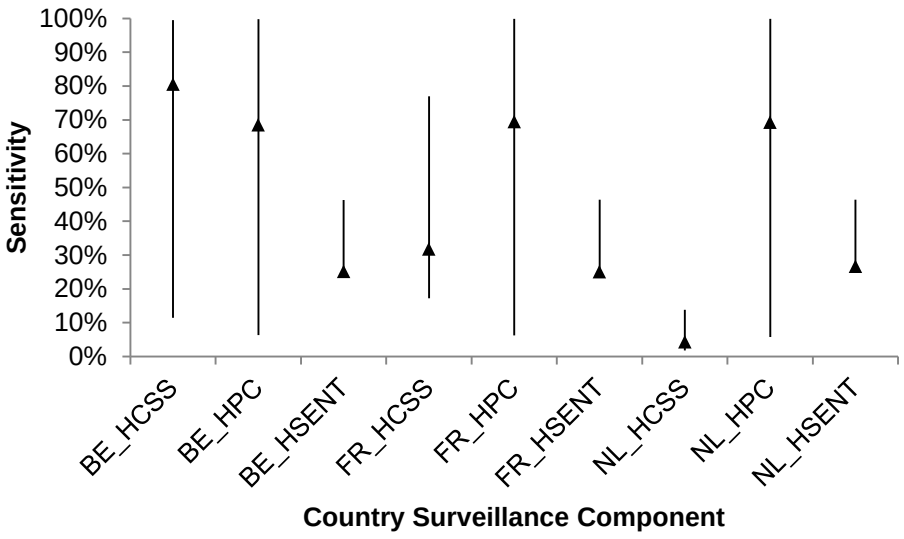


Figure 1: Simulated herd sensitivities (1st-50th-99th percentiles) in each component (CSS (Cross Sectional); SENT (Sentinel); PC (Passive Clinical)) in each country ((BE (Belgium); FR (France); NL (the Netherlands)).

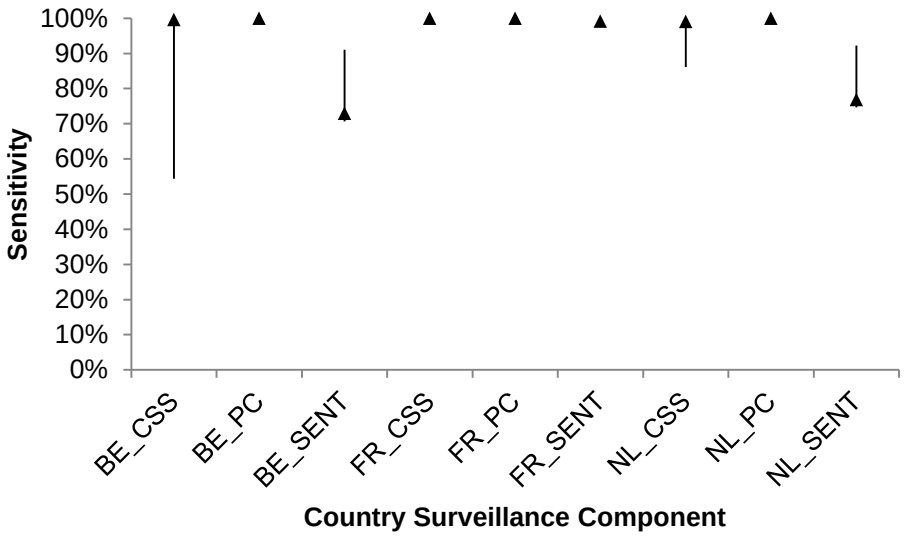


Figure 2: Simulated component sensitivities (1st-50th-99th percentiles) in each component (CSS (Cross Sectional); SENT (Sentinel); PC (Passive Clinical)) in each country ((BE (Belgium); FR (France); NL (the Netherlands)).

Though the mean herd sensitivities (depending on the diagnostic test used, number of cattle sampled within a herd and expected design prevalence) varied among the components considered, values were similar between countries, except for cross sectional surveys in the Netherlands (mainly due to the very small number of sampled cattle within herd). Component sensitivities were around 0.99 for most of the components except for the sentinel component (around 0.7).

Effectiveness of surveillance in terms of early detection

Table 5 displays the distribution of surveillance component sensitivity values obtained for early detection. Values are shown at herd (HSe) and component level (CSe). Two different simulations were done to account for the different within herd expected prevalence (2% (Sim 1) and 20% (Sim 2)).

Table 5: Mean (Minimum-Maximum) sensitivity values obtained at herd (HSe) and component level (CSe) for each component in Belgium and the Netherlands for each simulated scenario (2% (Sim1) and 20% (Sim2) simulated within herd prevalence).

Component level	***SIM 1_BE***	***SIM 1_NL***	***SIM 2_BE***	***SIM 2_NL***
Cross Sectional				
HSe	0.044(0.000-0.293)	0.002(0.000-	0.316(0.001-	0.019(0.000-
CSe	0.237(0.001-0.867)	0.176(0.001-	0.753(0.002-	0.729(0.002-
Sentinel				
HSe	0.007(0.000-0.027)	0.008(0.000-	0.072(0.001-	0.077(0.001-
CSe	0.038(0.0001-	0.043(0.001-	0.299(0.002-	0.328(0.003-
Passive Clinical				
HSe	0.159(0.001-0.706)	0.213(0.002-	0.678(0.007-	0.788(0.012-
CSe	0.999(0.375-1)	0.999(0.837-1)	0.999(0.995-1)	0.999(0.999-1)
Milk production data				
HSe	0(0-0)	0.082(0.001-	0(0-0)	0.494(0.002-
CSe	0(0-0)	0.985(0.035-1)	0(0-0)	0.999(0.525-1)
Reproductive performance				
HSe	0.005(0.000-0.018)	0.005(0.000-	0.049(0.001-	0.049(0.000-
CSe	0.578(0.005-0.971)	0.767(0.012-	0.977(0.090-1)	0.992(0.029-1)
Syndromic Mortality				
HSe	0(0-0)	0.002(0.000-	0(0-0)	0.015(0.001-
CSe	0(0-0)	0.256(0.001-0.641)	0(0-0)	0.867(0.017-0.999)

In general, sensitivities at herd and component level increased when the expected within herd design prevalence is set at 20%, with the exception of the component sensitivities of the passive clinical component and syndromic surveillance component based on milk production data.

In terms of early detection, it's mainly the passive clinical component and syndromic components that performed best (Table 5). In Belgium, using reproductive performance data was effective for early detection. No alerts were triggered using milk production and mortality data in Belgium; therefore CSe values were null for these components in Belgium. In the Netherlands, reproductive performance data and milk production data were effective for early detection as well as mortality data but the CSe of mortality was slightly lower compared to the other syndromic surveillance components.

Posterior probabilities of Freedom

Figure 4 displays the monthly posterior probabilities of freedom obtained in each month for each component in each country (using a within herd prevalence of 20%, probability of introduction per month of 0.5 constant throughout the year) and component sensitivities for the month in which the component was carried out.

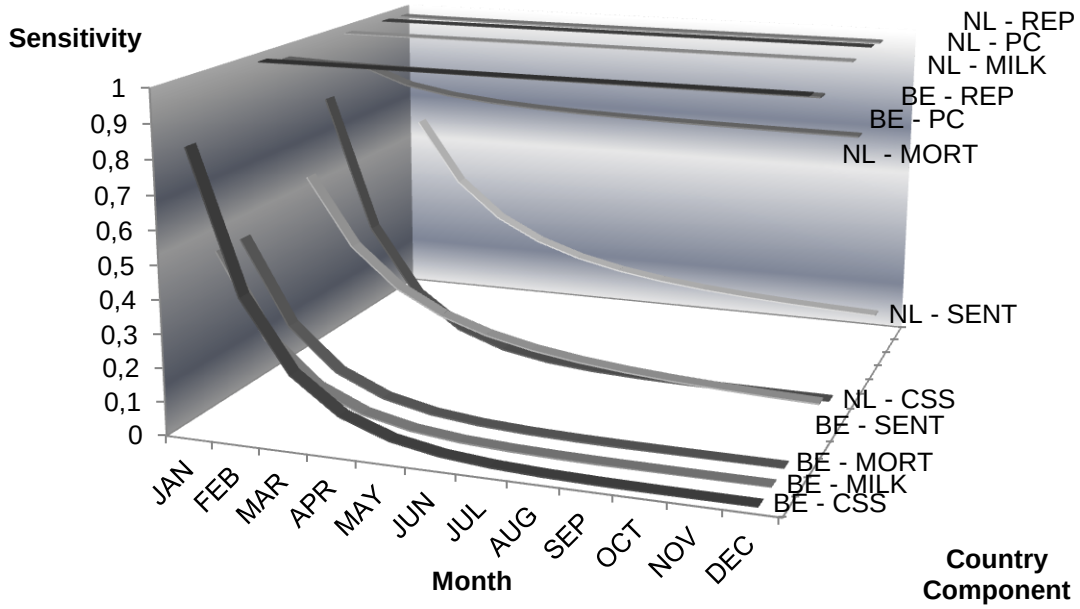


Figure 4: Monthly probability of freedom for each component (CSS (Cross Sectional); SENT(Sentinel); PC (Passive Clinical); MILK(Milk production data); REP(Reproductive performance data); MORT(Mortality)) in Belgium (BE) and the Netherlands (NL) given a constant probability of introduction (=0.5).

Assuming passive clinical surveillance is properly carried out, that component alone could provide sufficient guarantee towards freedom (CSe > 99%) in both countries. Syndromic surveillance based on reproductive performance data provided similar guarantees in both countries. Only in the Netherlands, milk production data provided sufficient guarantee of freedom. Despite the relatively good sensitivity of detection values for the remaining components, the limited monthly or daily sampling frequency and/or the limited number of herds surveyed were not enough to ensure freedom if the disease introduction probability was constant and equal to 0.5.

Cost efficiency

Table 6 displays the mean cost efficiency ratio, taking into account the mean component sensitivity (CSe) values for early detection in Belgium (assuming an expected within herd prevalence of 20%) and cost estimation, taking into account the number of animals and herds sampled.

Table 6: Mean sensitivity values obtained at component level (CSe) for each component in Belgium and cost estimation (in euros), taking into account the number of animals and herds sampled (na, nh).

Component	***Sampled herds***	***Sampled within herd***	***Cost total (in euros)***	***CSe***	***CSe/Cost***
Cross Sectional	344	40	146296	0.76	$0.52\%\ 10^{-3}$
Sentinel	260	15	74190	0.31	$0.41\%\ 10^{-3}$
Passive clinical	700	3	72410	0.99	$1.38\%\ 10^{-3}$
Milk production data	15	40	132635	0	$0.0\%10^{-10}$
Reproductive performance data	15	40	132635	0.97	$0.74\%\ 10^{-3}$
Mortality data	15	40	132635	0	$0.0\%10^{-10}$

For the purpose of early detection, total direct costs were lowest for sentinel and passive clinical surveillance. Passive clinical surveillance was most cost efficient, followed by syndromic surveillance based on reproductive performance data.

DISCUSSION

International standards more and more require surveillance to be fit for purpose and tailored according to the minimum guidelines and country needs. Moreover, effectiveness (FAO, 1999; OIE, 2014) and efficiency (Drewe et al., 2012; Häsler et al., 2011) are important parameters to consider while evaluating surveillance. Our first aim was to evaluate the effectiveness of each ongoing component for

substantiating freedom of disease, taking into account the minimum guidelines stated in the different European regulations, using Bluetongue surveillance as a case study. Using surveillance data of 2006-2007, the passive clinical surveillance components in Belgium, France and the Netherlands and cross sectional surveillance components in Belgium and the Netherlands were effective for substantiating freedom of infection after a whole year with component sensitivities values around 0.99, assuming a design prevalence of 2%. The sentinel component, mainly due to the small number of herds sampled, was of limited value with a CSe around 0.7. The herd sensitivities were slightly lower yet very much influenced by the number of animals sampled within a herd.

Our second aim was to determine effectiveness of the ongoing surveillance components for early detection as well as quantifying the added value of syndromic surveillance. Syndromic surveillance has gained increasing interest in recent years (Dorea, 2013; Madouasse et al., 2014). These studies have shown how syndromic surveillance could trigger alerts, but to date none published their effectiveness and cost efficiency in comparison to existing systems. In the present study we tried to tackle this issue and propose a method that could be used for this purpose. The approach elaborated in the present study combined existing methods, namely scenario tree models and epidemic curve models. A mathematical model, using epidemic curves obtained with Reed-Frost models, was developed for quantifying the effective probabilities of early detection for each surveillance component. This allowed comparison between components using a standardised approach. The effective probabilities of detection obtained revealed the subjective hypothesis that syndromic surveillance did not perform better than existing active surveillance components for sensitivity of early detection. Using this approach, passive clinical surveillance remained the best component in both countries. However over the year, syndromic surveillance based on routinely collected reproductive performance data in the Netherlands and Belgium and milk production data in the Netherlands could be of added value to declare freedom. No alerts were observed using mortality data in the Netherlands and Belgium and milk production data in Belgium, therefore sensitivity of these components were null. It could be argued that the reproductive ratio values used in the current study might change according to the disease under investigation.

However the stochasticity of the model took into account variability around the reproductive ratio and therefore provided reliable output that could be transposed to other epidemic diseases with similar transmission parameters to bluetongue. Despite the fact that the model provides interesting insight, relying on historical data for comparing surveillance components with a standardised approach and evaluating effectiveness of surveillance for similar diseases as Bluetongue, care must be taken in the interpretation of the output. Also, should a new disease emerge (with different clinical symptoms as Bluetongue) the outcome could change. If clinical signs are non-visible (subclinical infection/vaccination/natural immunity) or if repercussion on production data is not as evident the output is likely to change and these components might no longer be as effective. In the latter situation carrying out sentinel surveillance once suspicions are raised could definitely be effective. It must be noted that in the Netherlands during the second bluetongue epidemic wave that occurred in 2007, the first cases were found with the sentinel component (Santman-Berends et al., 2010). Also the efficacy of syndromic surveillance is strongly dependent on the alerts which are only triggered once sufficient amount of herds are affected by the disease. In addition, a main assumption is that the availability of data and regression lines and deviations of trends should be investigated on a daily basis. Therefore, it might be questionable in that respect whether syndromic surveillance will really enable early detection.

The EU legislation does not consider a herd effect for BT while estimating appropriate sample sizes. Therefore it can be expected that once the virus is present in a region it will rapidly spread to reach a within herd and between herd close to 100%. However, previous studies have shown that at the start of the epidemic the within herd prevalence was much lower (Méroc et al., 2008; van Schaik et al., 2008), also in a context of disease freedom or early detection, it seems reasonable to consider very low within herd prevalence. This is the reason for simulating within herd expected prevalence of 2% and 20% and measure the impact on the herd and component sensitivity. The obtained sensitivities varied greatly according to the chosen prevalence, revealing the importance of considering correct expected within herd and between herd prevalence when setting up or evaluating surveillance. For the rest of the study we used a between herd prevalence of 2% and within herd

prevalence of 20%, as we considered that would a herd be infected with Bluetongue due the vector borne disease nature of the epidemic, it would be reasonable to expect a 20% within herd prevalence.

Passive clinical surveillance and syndromic surveillance based on reproductive performance data in both countries, as well as syndromic surveillance based on milk production data in the Netherlands, provided best levels of the posterior probability of freedom. The daily observations generated by these components as well as the high number of herds included, contributed to this high level of guarantee, together with the ability of generating alarms. However, this capability, again, strongly relies on the given fact that the considered vector-borne disease will show visible clinical signs or reproductive effects. If the considered disease would not show significant clinical signs, these components would probably not be appropriate.

From the cost point of view, the total direct cost was lower with sentinel and passive clinical surveillance. Yearly cross sectional surveys have the limitation that it is not repeated within a year therefore hampering early detection. However the aim of a yearly cross sectional survey is to provide information on the disease status (freedom or prevalence) in a country and not early detection, assuming sufficient cattle within herd are sampled. Passive clinical surveillance offered not only good sensitivities of detection, was also most cost efficient providing the assumed direct cost. Syndromic surveillance is also cost efficient. Should we account for indirect cost, it might be considerably less cost efficient. Indeed, Hanon et al., (2009) revealed the total costs to be over several millions of euros. Besides the fact that, as discussed above, the disease under consideration must have a visible impact on the production parameters considered for syndromic surveillance, lack of specificity together with poor positive predictive value, due to the false alarms and the difficulty of differentiating a baseline trend behind the background noise, constitute another major limitations of syndromic surveillance (Hope et al., 2006). The poor specificity might lead to extra cost linked to confirmation testing generated in order to rule out the suspicious cases. In the present study we considered 15 false alarms per year. Despite this, syndromic surveillance offers promising diagnostics of deviations from trends, and measuring the impact of disease, using readily available data (Madouasse et al., 2014). In a context of globalisation, where emerging diseases

spread and cross borders fairly rapidly and when disease awareness is not always present, syndromic surveillance definitely has an important role for surveillance.

Data availability and quality is often neglected in surveillance evaluation, thereby hampering inference that can be made of surveillance simulations. The current study corroborated the essential preliminary requisite of data in order to guarantee effectiveness while conducting surveillance evaluation. Indeed the limited availability of data for France illustrates the importance and impact of lack of national data and how it affects epidemiological work. Also, work carried out and experience within this transboundary project provided useful insight about the critical hurdles and solutions one should consider while modelling disease across borders.

In conclusion, the standardised approach developed in the present study, can be considered as a powerful tool for quantifying effectiveness in early detection, and no matter what disease is concerned, providing reliable data is used to fit the models. We would recommend from the present study to use passive clinical and syndromic surveillance on a routine base for early detection and impact estimation, in addition, to implement existing traditional surveillance methods, such as repeated cross sectional survey surveillance, once suspicions are raised. This would probably be the most effective and most cost efficient option.

REFERENCES

de Koeijer A, Jan Boender G, Nodelijk G, Staubach C, Meroc E, RW Elbers A. Quantitative analysis of transmission parameters for bluetongue virus serotype 8 in Western Europe in 2006. Vet Res. 2011; 42(1): 53.

Dohoo, I., Martin, W., Stryhn, H. (Eds.), 2009. Targeted sampling. In: Veterinary Epidemiologic Research. VER Inc, Canada., pp. 41-42.

Dórea FC, McEwen BJ, McNab WB, Revie CW, Sanchez J, 2013. Syndromic surveillance using veterinary laboratory data: data pre-processing and algorithm performance evaluation. J R Soc Interface. 10(83): 20130114.

Drewe JA, Hoinville LJ, Cook AJ, Floyd T, Stärk KD, 2012. Evaluation of animal and public health surveillance systems: a systematic review. Epidemiol Infect. 140(4):575-90.

European Commission (EC), 2012. Commission Regulation (EC) No. 1266/2007 on implementing rules for Council Directive 2000/75/EC as regards the control,

monitoring, surveillance and restrictions on movements of certain animals of susceptible species in relation to bluetongue. Official Journal of the European Union L141, 7–13.

Food and Agriculture Organisation (FAO), 1999, Manual on Livestock Disease Surveillance and Information Systems. URL (Last update date: no date) http://www.fao.org/docrep/004/x3331e/x3331e01.htm (Consulted date: 14/07/2014).

Frössling J, Wahlström H, Agren EC, Cameron A, Lindberg A, Sternberg Lewerin S., 2013. Surveillance system sensitivities and probability of freedom from Mycobacterium avium subsp. paratuberculosis infection in Swedish cattle. Prev Vet Med. 108(1):47-62.

Hanon JB , Uyttenhoef A, Fecher-bourgeois F, Kirschvink N , Aubruge E ,Duquesne B, Saegerman C., 2009. Estimation des pertes économiques attribuées à la fièvre catarrhale ovine (sérotype 8) dans le sud de la Belgique durant la période 2006-2007. Epidémiol et santé ani, 56 :187-195 .

Häsler B, Howe KS, Stärk KDC, 2011. Conceptualising the technical relationship of animal disease surveillance to intervention and mitigation as a basis for economic analysis. BMC Health Serv Res. 11: 225.

Hope K, Durrheim DN, d'Espaignet ET, Dalton C. Syndromic surveillance: is it a useful tool for local outbreak detection? J Epidemiol Community Health. May 2006; 60(5): 374–375.

Knight-Jones TJ, Hauser R, Matthes D, Stärk KD, 2010; Evaluation of effectiveness and efficiency of wild bird surveillance for avian influenza. Vet Res. 41(4):50.

Madouasse A, Marceau A, Lehébel A, Brouwer-Middelesch H, van Schaik G, Van der Stede Y, Fourichon C, 2013.Evaluation of a Continuous Indicator for Syndromic Surveillance through Simulation. Application to Vector Borne Disease Emergence Detection in Cattle Using Milk Yield. PLoS One. ; 8(9): e73726.

Martin, P.A., Cameron, A.R., Greiner, M., 2007. Demonstrating freedom from disease using multiple complex data sources 1: a new methodology based on scenario trees. Prev. Vet. Med. 79: 71-97.

Méroc E, Faes C, Herr C, Staubach C, Verheyden B, Vanbinst T, Vandenbussche F, Hooyberghs J, Aerts M, De Clercq K, Mintiens K, 2008. Establishing the spread of bluetongue virus at the end of the 2006 epidemic in Belgium Vet Microbiol. 131(1-2):133-44.

Saegerman, C., Mellor, P., Uyttenhoef, A., Hanon, J.B., Kirschvink, N., Haubruge, E., Delcroix, P., Houtain, J.Y., Pourquier, P., Vandenbussche, F., Verheyden, B., De Clercq, K., Czaplicki, G., 2010. The most likely time and place of introduction of BTV8 into Belgian ruminants. PLoS One 5(2), e9405.

Santman-Berends IM, Bartels CJ, van Schaik G, Stegeman JA, Vellema P, 2010. The increase in seroprevalence of bluetongue virus (BTV) serotype 8 infections and associated risk factors in Dutch dairy herds, in 2007. Vet Microbiol. 142(3-4):268-275.

Santman-Berends IM, Stegeman JA, Vellema P, van Schaik G, 2013. Estimation of the reproduction ratio (R(0)) of bluetongue based on serological field data and comparison with other BTV transmission models. Prev Vet Med. 108(4):276-284.

Vandenbussche F, Vanbinst T, Verheyden B, Van Dessel W, Demeestere L, Houdart P, Bertels G, Praet N, Berkvens D, Mintiens K, Goris N, De Clercq K. Evaluation of antibody-ELISA and real-time RT-PCR for the diagnosis and profiling of bluetongue virus serotype 8 during the epidemic in Belgium in 2006. Vet Microbiol. 2008 May 25;129(1-2):15-27.

van Schaik G, Berends IM, van Langen H, Elbers AR, 2008. Seroprevalence of bluetongue serotype 8 in cattle in the Netherlands in spring 2007, and its consequences. Vet Rec. 163(15):441-444.

Veldhuis A, Brouwer-Middelesch H, Marceau A, Madouasse A, Van der Stede Y, Fourichon C, Welby S, Wever P, van Schaik G. Application of syndromic surveillance with routinely collected cattle herd production data for early detection of outbreaks of Bluetongue and Schmallenbergvirus. Submitted

Welby S, Méroc E, Faes C, De Clercq K, Hooyberghs J, Mintiens K, Van der Stede Y, 2013. Bluetongue surveillance system in Belgium: a stochastic evaluation of its risk-based approach effectiveness. Prev Vet Med. 112(1-2):48-57.

Wilson AJ, Mellor PS, 2009. Bluetongue in Europe: past, present and future. Philos Trans R Soc Lond B Biol Sci. 364(1530):2669-2681.

World Organisation for Animal Health (OIE), 2014. Guidelines for Animal Disease Control. URL (Last update date: 15/05/2014): http://www.oie.int/fileadmin/Home/eng/Our_scientific_expertise/docs/pdf/A_Guidelines_for_Animal_Disease_Control_final.pdf (Consulted date: 14/07/2014).

Zientara S, Sánchez-Vizcaíno JM, 2013. Control of bluetongue in Europe. Vet Microbiol. 165(1-2):33-37.

Chapter VIII: GENERAL DISCUSSION

The overarching purpose of this thesis was to define a comprehensive and objective quantitative approach for the evaluation of animal disease surveillance systems. More specifically, different applications of a published method (output based methods) were explored to study the different possibilities offered by these tools and define their limits and constraints. By testing surveillance systems performance against stated objectives and targets, taking into account local risk factors such as heterogeneity of husbandry practices and epidemiological situations, as well as the needs and priorities expressed by the stakeholders involved, results of this thesis indicate what might be useful to guide prioritisation in the surveillance of animal health hazards and risks.

This thesis has shown how simulation models can contribute to document the efficacy of surveillance but also guide decision towards most performing surveillance alternatives in line with international recommendations regarding transparency (EC, 2016).

RISK BASED SURVEILLANCE TO IMPROVE EFFICACY OF SURVEILLANCE COMPONENTS

The probability that an animal or herd tested negative is truly free of disease may vary from one population to another. Biological variation, such as stage of infection or farm past disease status, can provide explanations for differences but so can non-biological causes, such as social and political drivers between similar eradication and surveillance programmes as applied in different countries. In addition, similar surveillance and control measures may not have the same impact or value when applied under different epidemiological circumstances. The existence of a range of epidemiological factors operating at a local level (such as farm husbandry practice, on farm movements and biosecurity levels, for example) generates complex situations that influence the effectiveness of both surveillance and control activities. The following sections will discuss these issues and indicate ways of improving surveillance.

Risk based surveillance

Chapter III, IV, V and VI showed how output based and risk based approaches to surveillance have the potential to be more flexible and eventually provide higher

confidence towards the stated objective of surveillance by taking into account country specific risk factors. Targeted surveillance combines the concepts of hazard specific surveillance (targeted to specific pathogensand of risk based sampling (targeted to groups with higher risk of infection) as clearly shown in chapter III in which samples for AI were reallocated to regions at a higher risk when compared with others. In chapter IV it has been made clear that the Brucellosis surveillance programme should make more efforts to increase the number of samples obtained from the abortion (protocol) component. If target surveillance is applied, it will ultimately enable better detection and an overall reduction of total sample size by focussing surveillance activities in subpopulations most at risk of infection; thereby increasing the probability of detection and reducing costs, as corroborated by others (Bisdorff et al., 2016; Hoinville, 2013; Knight-Jones, 2010; Rutten et al., 2012).

However, in order to apply risk based surveillance, identification of risk factors is a crucial step to determine these different probabilities of detection/infection but it might also be useful to adapt control or intervention measures in line with their risk profile. Herds or animals can be scored on the basis of known risk factors estimated from literature (chapter III), by stakeholders needs and priorities (chapter IV) or from historical data (chapter V, VI). The advantage of the latter is seen when country specific data can be used to model the probability of a disease occurrence given known and recorded country specific risk factors (chapter V, VI).

From the different studies outlined in the previous chapters, the key elements to consider in light of their differential risk of infection or infecting profile when scoring herds or animals surveillance & intervention programs can be summarised as follows: centralisation and recording of data regarding herd localisation, herd type (dairy, meat, hobby), ages, density or size, herd past disease status, movement records (making use of available national animal identification registries (SANITEL), trade (SANITEL, TRACES)), poor biosecurity measures, on farm delivery of raw dairy products, frequency of testing schemes and results (positive as well as negative or doubtful), herds exempted of testing schemes or not tested for long periods (i.e. because considered as dead-end reservoir). At the time of the study presented in chapter III, relative risks were estimated on the basis of literature and neighbouring country data. Subsequently, a research project (Flutree) (van den Berg et al., 2010),

was initiated to better understand some specific characteristics of the Belgian poultry farms and wild bird populations. During this research project, it was found that preventing outdoor feeding could largely reduce the probability of wild bird incursions. Among professional poultry holdings, those with outdoor access and several bird species raised on the same farm had a higher probability of infection than others. Following these findings it could be recommended to intensify surveillance in farms with less biosecurity as corroborated by others (Caron et al., 2014; Conan et al., 2013; EFSA, 2017; Ferrer et al.,2014). Compulsory testing at farms that sells cattle rather than at farms of arrival could potentially avoid consequences of disease spread, but separate barns for purchased cattle should be available in farms and newly introduced animals should be kept under observation for a while before mixing them with other animals especially on intensive production farms where disease introduction could affect very rapidly the whole herd. Veterinarians should inform their clients about correct biosecurity measures. It is comforting to notice that the same risk factors identified in the Belgian context have been recognised in other countries by others. For instance, for bovine tuberculosis surveillance, looking into Belgian historical cases data, past tuberculosis status of the farm, movement rate and herd size were clearly identified as risk factors for infection and were considered as a necessary surveillance focus as demonstrated by others too (Adkin et al., 2016;, Bisdorff et al., 2016; More et al., 2015). The risk based surveillance (essentially focussing surveillance on testing of abortions and herds where abortions are expected to have occurred but were not notified) showed that a 25-fold reduction of the number of samples was feasible and still maintain the desired confidence level regarding the country's disease status. These implemented alternatives revealed themselves to be pivotal for increase confidence in bovine brucellosis surveillance sensitivity but also for its negative predictive value and therefore useful for both objectives (increasing probability of freedom and enable early detection) taking into account both risks probability of infection (upstream surveillance) and risk consequence of infection (downstream surveillance). Following this study the notification of abortions increased (from 4056 in 2008 to 11836 in 2015 (national yearly reports)). This abortion notification led to the detection of B. abortus in 2010 and 2012 (OIE, 2012). The intense epidemiological investigation enabled rapid

detection and limitation of the infection by massive testing and elimination of positive reactors. This illustrates the positive impact of campaigns that were initiated using incentives to encourage farmers to report abortions resulting in an increasing number of reported cases. Since then, similar strategies have been identified in neighbouring countries (Bronner et al., 2013, 2014, 2015).

Thus, the Belgian field experience of setting up risk based surveillance confirms the theoretical concept that targeting sampling using risk consequence to increase surveillance efficiency can be implicitly considered as a way to increase case detection. Targeting surveillance in risk groups where probability of occurrence is high, allows increasing surveillance efficacy to demonstrate freedom (Cameron, 2012).

Perception and prioritisation of risk as well as acceptability towards the stated objective of surveillance

The interplay between international and national policies, local risk factors and different actors in the disease surveillance highlight the complexity of surveillance. Indeed, risk perception and priorities may differ significantly between stakeholders and policy makers, emphasising the essential role of communication between all partners.

Chapter IV first considered stakeholder's opinions regarding the different relative risks and simulated different scenarios to determine the impact of different surveillance strategies for Brucellosis surveillance. After which, results of the simulated scenarios were shown to the same stakeholders and policy makers, whom then had the opportunity to review their risk estimation value. In the absence of data or when budgetary constraints limit field investigation, this approach showed itself of high value to guide decision. Following this two step process, a consensus was found on optimal risk parameters closest to real life. Bronner et al. (2014, 2015) confirm the impact of stakeholder's involvement in surveillance such as for bovine brucellosis surveillance performance in France.

Decision makers facing a situation expect guidance from experts whilst experts often answer "it depends on the situation". Because a clear question or a clear model to answer the question is not always available, one must tailor models and adapt the questions to needs. Simulation models and sensitivity analysis appear to be useful

tools that can measure the impact of over or under estimating situations such as experienced in this two-step EKE method.

Participative and social epidemiology has increasingly gained interest and provides useful qualitative tools to ease communication and clarify expectations from surveillance systems (Calba et al., 2015; Cowie et al, 2015; Stark & Hasler, 2015). Qualitative approaches for surveillance evaluation such as the OASIS model (Calba et al., 2016; Hendrikx et al., 2011) or SERVAL model (Drewe et al., 2015) methods, measure the acceptability of surveillance. RESET (Rules, Education, Social pressure and Economic stimuli Tools) framework focuses on elements of farmers' mind sets that are influential in disease control, including perceived threats, and perceived efficacy of preventative measures (Jansen et al., 2012). Initially developed for handling mastitis on farm these tools could be extended to disease control and surveillance.

Identifying incentives to encourage farmers and veterinarians in complying with surveillance and acceptability of surveillance has shown to be beneficial to Belgium. Indeed ARSIA found in 2013 a lower reported rate of abortion in contrast to 2009, 2010, 2011, 2012, where a constant increase was observed explained mainly by the incentives (free testing of abortions for several disease) introduced to comply with the benchmark set up by ERASURV in order to meet the required surveillance systems sensitivity (chapter IV). Following the unexpected success of these incentives the budget for abortion surveillance exploded in to the extent that authorities alone could not bear the cost of this surveillance and ARSIA accepted to fund additional analysis within the abortion protocol in order to guarantee the compliance of farmers and vets with the surveillance system. The decrease of the Schmallenberg epidemic wave in 2012 together with the resolution of the brucellosis outbreaks could also explain this slight decrease in abortion reports observed in 2012-2013. However this rapidly shifted in 2014 in such a way that it could be concluded that abortion reporting incentives and target sampling of holdings that not report an abortion had enabled an increase of 244% compared to 2008 in Wallonia. At national level, notification of abortion increased from 4056 in 2008 to 11836 in 2015. Despite this steady increase in abortion reporting from 2009 up to present, underreporting is still suspected and this will have an impact on the negative predictive value of surveillance. This is why

brucellosis surveillance does not only rely on abortions notification in Belgium. In contrast to other European Member States officially free of brucellosis, in Belgium we have re-implemented bulk milk testing as well as during the winter target sampling of herds that did not notify an abortion, but yet are suspected to have had an abortion (based on previous year's annual birth rates and submission of low weight carcasses sent to rendering plant as well as diagnostics lab results). From our experience during this work, a hazard analysis critical control plan (HACCP) could be considered simultaneously to a the output based methods using scenario trees, to identify, characterise and manage the critical points, gaps and needs in surveillance. This process should be updated constantly in the light of new evidence. Ensuring a comprehensive approach while evaluating surveillance systems would probably increase compliance and, ultimately, increase effectiveness and efficiency of surveillance but also, within time ensure sustainability too. Understanding the context and drivers of surveillance is critical in designing successful interventions actions. These intervention actions will only be effective if key players in the system accept them.

Combining different sources of evidence for different surveillance objectives

In contrast to Chapter III, which only investigated one component of surveillance (active monitoring program), Chapters IV, V, VI and VII combined information obtained from different data sources and components to strengthen the confidence in probability of freedom and/or case detection. Pursuing with this same notion, it can even be determined what component can eventually be disregarded for not bringing sufficient information, or for providing poor quality information (Chapter V). Indeed, pooling all components together may provide more guarantees than needed for a given objective, i.e. to claim freedom. However, it is important to measure the efficacy of each component in relation to the desired goal or attribute one opts to assess when conducting surveillance (Drewe et al., 2012; Hoinville et al., 2013). Therefore, it may be useful to maintain some components, such as purchase testing or abortion testing (Chapter IV, V) to maintain disease awareness and early detection. However there will be population overlap as the same population might be looked at in several surveillance components, meaning that surveillance components are not totally independent. The overall surveillance system sensitivity will be increased if several surveillance components are in place, in such a way that it might lead to an overestimation of the negative predictive value of your surveillance system if this overlap is not taken into account.

Methods proposed by Martin et al. (2007) enable to take into account this overlap by adjusting the effective probability of infection within the current component by posterior probability of infection gained from the previous surveillance component. One could measure the population overlap (=animals sampled in several components and remove them from the calculation of the given component if one has already counted them in the previous component). It might be useful to consider each component separately as it may be desirable to maintain each different component even though the total surveillance system sensitivity will be reached with or without each of them, as each component might have different purposes.

In this respect, Chapter VII looked at the cost efficiency and effectiveness ratios of conventional surveillance components compared with emerging surveillance techniques such as syndromic surveillance in relation to the stated objective. It was discovered that although the cheapest alternatives were sentinel surveillance and passive clinical surveillance for substantiating freedom, passive clinical and syndromic surveillance were, in contrast, were most cost efficient for early detection and provided best posterior probabilities of freedom , by accumulating evidence dayly over time. In addition, syndromic surveillance could be useful to carry out impact analysis or increase specificity once disease is detected. As suggested by Drewe et al. (2012), it is useful to consider more than two attributes to evaluate surveillance performance in addition to clearly stating the objective. Our findings suggest that sensitivity, negative predictive value, specificity, timeliness and data quality are fundamentally important to measure and that they could all be reduced to gauging efficacy and efficiency of surveillance system performance for possibly different stated objectives. In addition, performance of surveillance systems should be expanded to all sectors (from field to food and/or human cases) while evaluating surveillance measures but also all different possible data sources contributing to surveillance, providing in such a way an integrated approach to optimise surveillance and sampling (Häsler et al., 2016; Martins et al., 2016). Acceptability should also be assessed from a public health and social point of view for prioritisation of disease surveillance or eradication measures. International requirements, but also animal welfare issues, economical consequences, zoonotic potentialof risks should be accounted for in the evaluation of the cost of surveillance.

ESSENTIAL ASSUMPTIONS IMPACTING SURVEILLANCE PERFORMANCE

Design prevalence

Design prevalence is part of the design of the model and is not related to actual prevalence; it is set for the purpose of drawing conclusions about the effectiveness of the surveillance component against an agreed standard. Design prevalence will enable to test how confident one can be in a surveillance system given no test positive animals are found. It It could be argued that the lower the deisign prevalence the better the 'early detection' in the sense that waiting for disease prevalence to reach a higher level does not coincide with 'early detection'. However, to measure the surveillance system's reliability in early detection one would rather recommend to use the approach developed in chapter VII, using reed frost models outputs to feed the scenario tree inputs and measure the performance of surveillance in terms of early detection. Although most official regulations do mention requirements in terms of maximum tolerable herd prevalence, very few mention tolerable within herd or animal prevalence. Examples of this are the Council's directive 64/432/EEC regarding Brucellosis and Tuberculosis, where nothing is mentioned regarding expected animal prevalence. This is in contrast with the legislation on avian influenza (2007/268/EU), where design prevalence requirements for between herd (5%) and within herd (30%) are clearly mentioned. Despite the fact that these official requirements can be questioned to be rightly selected or not, surveillance programmes often rely on such values for their design and hence will influence the results of these programmes. Either one can assume 100% herd sensitivity, implying the right tests are used and the right number of animals per herd are sampled depending on the herd size, or one can assume that if disease occurs within a herd the whole herd will shortly be infected to reach 100% within herd prevalence, thus maximising herd sensitivity near 100%. In practice, and especially for diseases such as Brucellosis and Tuberculosis, neither of the assumptions (all animals tested with perfect test sensitivity and specificity or 100% homogeneous prevalence within the herd) can be verified and thus, 100% herd sensitivity is rarely met. In Belgium, during the last outbreaks of Brucellosis, within herd prevalence was rarely above 2-3% and in some very rare cases it was no higher than 12% (AFSCA outbreak data). In

Chapter V, because again nothing is mentioned regarding minimal expected within herd prevalence for bovine tuberculosis, it was assumed that 17% was reasonable to expect, based on historical data investigation data (AFSCA outbreak data). This underlines the importance of adapting sample size accordingly (Chapter VII). Also herds tend to get larger and larger, triggering the need to adapt within herd sample size accordingly (Eurostat, 2013).

According tochapter VII, simulations carried out regarding bluetongue also clearly proved the importance and impact of assumptions regarding expected within herd prevalence. Ensuring disease freedom just below the official design prevalence, which satisfies international standards, does not explicitly mean one is totally free. Also, a country with weak surveillance may achieve the minimum requirements more easily than a country having a more stringent surveillance (Stärk and Häsler, 2015). This could be the paradox of official minimum requirements and freedom when designing surveillance. Maybe an alternative could be to allow each EU Member State to demonstrate freedom with a matching maximum tolerable prevalence. This matching prevalence would have to be proven. This would clearly be of value to countries with highly heterogeneous herd sizes (i.e., very small and very large) where it will be complex to obtain 100% herd sensitivity. It must therefore be argued whether sensitivity of surveillance assessed at the prescribed design prevalence ensures an acceptable negative predictive value of surveillance (Bisdorff et al., 2016; Cameron 2012).

Diagnostic tools: sensitivity, specificity and timeliness

In contrast to public health essentially relying on passive notification, animal health surveillance relies on active surveillance involving the selection and testing of animals. Our different chapters illustrated the great range of variability in different outputs arising from uncertain test characteristics, underlining thereby the strong emphasis that should be placed on this parameter. More specifically, Chapter V showed that for bovine tuberculosis the impact of meat visual inspection sensitivity on the confidence level of our posterior probability of freedom. Our results corroborate previous findings and the repercussion on surveillance negative predictive value suggested by others (Calvo-Artavia 2013; Foddai et al., 2015; Stärk et al., 2014). Besides the diagnostic test sensitivity, the completeness and coverage

of slaughterhouse surveillance will naturally impact the performance of surveillance as well (Jajosky and Groseclose , 2004; Reijn et al., 2011; Welby et al., 2012, 2015). However, specificity, often disregarded in these methods as considered to be 100% since it is assumed that any positive result in surveillance will be investigated until confirmed negative), merits to be considered as well in certain circumstances. Few publications in animal health surveillance highlight the impact of this assumption and often consider this same assumption as an asset. However, in a context of budget restriction and moving towards more cost efficient surveillance, we believe this is a major assumption that cannot be considered as an asset while considering the performance of surveillance. For diseases such as bovine tuberculosis, specificity plays a major role on cost efficiency of surveillance (Welby et al., 2015). At farm level this might seem of insignificant impact, however in a surveillance component processing over 345,000 cattle (as purchase testing in Belgium), when specificity is less than 100 per cent, this will have a major impact. Indeed, this means, a proportion of animals not infected will have a (false) positive reaction to the test, leading to further confirmation test or slaughter. In some cases as for bovine tuberculosis, a false positive test result will initiate or prolong the suspended tuberculosis free status, meaning in Belgium no more movements or on farm delivery of dairy products, which indirectly increases the costs of eradication campaigns to farmers and government and undermines stakeholder confidence (Goodchild et al., 2015; Nielsen et al., 2015). Imperfect specificity cannot be neglected and this of greater challenge when prevalence is very low (More and Good, 2015). Indeed, the cost of errors when prevalence reaches very low values, near zero (as in disease free situation), will be mainly driven by false positives reactions generated by a lack of specificity. While evaluating surveillance performance, these false positives reactions costs can neither be ignored nor neglected as this will generate extra costs. One way to take the specificity into account would be to consider it in the denominator of the cost efficiency ratio equation (numerator being sensitivity) and adapt control measures in case of suspicion proportional to the risk (Welby et al., 2015). Similarly, in chapter VII, cost of false alarms in syndromic surveillance showed that specificity should not be neglected. Vial et al. (2015) suggest, that in order to be cost efficient, false alarms should not be over 2 and at 6 monthly intervals, and they

also suggest to increase specificity with the use of a multivariate approach while using syndromic surveillance (not only relying on slaughterhouse data or mortality data but rely on milk production data and/or fertility data too), as corroborated in chapter VII. The same authors suggest that timeliness linked to processing and analyses should not be neglected as this adds to the cost as well (Vial et al., 2015). In chapter VII, timeliness was assessed using Reed-Frost models and the sensitivity of the diagnostic process was adjusted accordingly. To facilitate future comparisons of reporting timeliness across components and countries we also would recommend to include an explicit description of the surveillance system reporting process and interval. Additionally, surveillance information systems must support the collection of appropriate reference dates to allow for the assessment of the timeliness of specific surveillance processes.

Most EU countries rely on passive surveillance components as diagnostic process for case detection and confidence level in disease freedom. Impact of disease awareness or total number of animals inspected during passive clinical surveillance was investigated in Chapter VI, and the much larger confidence interval around the obtained output further confirmed the huge variability and uncertainty if one was to rely on this component alone. Fear of ethical or economic repercussions or repressive measures triggers underreporting. This apparent poor disease awareness amongst farmers and or vets in turn hampers effective surveillance as shown in chapter VI and VII and underlined by others (Bisdorff et al., 2016; Bronner et al., 2014, 2015). Representativeness as well as timeliness of passive surveillance remains a challenging issue for public health in general (Jajosky and Groseclose, 2004; Reijn et al., 2011). Despite the yearly steady increase in abortion notifications, a study carried out by ERASURV (Welby et al., 2014) corroborates findings of ARSIA (ARSIA, 2014) and French colleagues (Bronner et al., 2013) regarding suspicious under reporting of abortions. Rendering plant data (carcasses less than 10 kg or between 10-25 kg) and yearly birth rates for each active cattle farm in Belgium in 2012-2013 enabled ERASURV to conclude that only 36% of the herds that submitted low weight carcasses to RENDAC reported and submitted an abortion for further analysis in 2013. In other words, 64% of the herds did not report any abortion and yet submitted a low weight foetal carcass to RENDAC. These figures are similar to data

reported by our French colleagues (Bronner et al., 2013, 2014). There are some of the reasons that could explain this: foeti of poor quality and/or discovered too late in the field and contaminated to such an extent that is impossible to be analysed. In addition, 113 herds submitted low weight carcasses to RENDAC in 2013 and did not report any births neither did they notify abortions that could underline a fertility problem in those herds. This highlights that approximately 50% of abortions are not reported, despite all efforts by the competent authority to simulate and compensate the reporting and analysis of any abortion. Methods we would recommend to improve reporting of abortions are: i) Win-Win approach to surveillance (i.e. abortion protocol incentive: panel of test free of charge for the voluntary submission of abortive tissue samples); ii) Winter target sampling of herds not reporting abortion but suspected to have had an abortion in the past three years (based on birth rates, population data, rendering plant data, abortion diagnostic data); iii) Information campaigns for farmers and vets (conference, debates, flyers, mails); iv) Using tools at herd level to assess herd health; v) Increase compliance & acceptability (i.e. information campaigns, by providing incentives or reduce restriction measures applied to suspicious farms, avoid using tests with false positives reactions that could undermine the farmers believe in surveillance). Applying system thinking (conditional loops and behavioural over time plots) to veterinary epidemiology could help to better understand and increase disease awareness.

Input validation versus output validation

As the need for better quality control continues and standardisation is introduced in an emerging diversity of surveillance systems, validation surveillance systems evaluation is an essential step. Although sensitivity analysis using internal validation techniques provides useful insight, it will be strongly influenced and driven by the variability and uncertainty around the given inputs. Chapters VI and VII tested assumptions by changing one parameter while keeping all others unchanged. Through these studies, validation provided useful insight into the robustness of results and models as well as the magnitude of possible over and under estimation of outputs. However, besides quantitative validation, qualitative validation should be assessed too. Validation can be made in terms of effectiveness, using different attributes described to assess it, such as compliance to a surveillance system by

animal industry partners concerned. Where ante mortem and post mortem diagnostic tools have neither perfect sensitivity nor specificity as seen above this can lead to major impacts. In some cases, such as in bovine tuberculosis diagnostic testing schemes, clear discrepancies may be observed, leading to a wide variability of reported values (Bezos et a., 2014; Casal et al., 2015; Schiller et al., 2010). Internal sensitivity analyses carried out as in chapter V clearly corroborates this and illustrated the impact of imperfect diagnostic assays have on a larger scale. But external validation carried out as in chapter V, also revealed what was more fitting in a Belgian context. Exploring the 7,403,826 cattle movement history records that were obtained for the 2,678,020 cattle and 36,059 herds still active in 2009, coupled with historical bovine tuberculosis outbreak data, provided an insight into the main risk factors for bovine tuberculosis infections in Belgium (Chapter V). In addition, validation of component's performance was investigated, looking at historical data and modelling the detection odds of each component, using a generalised equation model. Results of the validation models suggested that despite cattle purchases identified as a significant risk factor, purchase testing with the single intradermal tuberculin test was non-efficient for the detection of outbreaks in Belgium, raising serious doubts on the trust one could place on Belgium’s official and bacteriological freedom status, since one of the main pillars (after slaughterhouse surveillance) of bovine tuberculosis surveillance in Belgium is purchase testing. It also highlighted the importance of recording data at an individual level in a centralised data base, currently not done for tuberculin tests carried out in Belgium. In addition the results showed the importance of risk based slaughterhouse surveillance to increase total sensitivity as demonstrated in other studies (Calvo-Artavia et al., 2012; Calvo-Artavia 2013; EFSA, 2013; Foddai et al., 2015; Stärk et al., 2014). The low sensitivity of intradermal skin test can be counteracted by larger sample sizes and this test has been reported to be a valid test at herd level (Schiller et al., 2010). However, despite the factual increased sample size, this test has shown its limits in the Belgian field. A benchmarking study ,carried out by ERASURV, showed that the amount of false positive reactions reported with this test, when used during purchase testing, was still worryingly low in comparison with the false positive reactor rate expected and observed in other circumstance (ie testing of dairy cattle with on farm delivery)

(Welby et al., 2015). This study revealed similar findings regarding slaughterhouse surveillance, thereby highlighting the major importance of awareness for the performance of these surveillance components (Welby et al., 2015). In addition, we would recommend addressing timeliness of detection when evaluating surveillance diagnostics, in order to enable fair comparison between countries and components, i.e. using Reed-Frost models (Chapter VII).

PERFORMANCE OF SURVEILLANCE: ACROSS COUNTRIES AND EMERGING DISEASES

Cross border surveillance and emerging or re-emerging disease

In 2006, an unexpected vector borne disease, named bluetongue serotype 8, emerged in three countries in Northern Europe, Belgium, Germany and The Netherlands. It had devastating effects on both trade and productivity. In order to regain an official freedom status and re-establish safe trade, Belgium had to demonstrate the efficacy of its surveillance to EU authorities, in both freedom of disease and early detection.

Chapter VII specifically explored different surveillance components across countries and their capacity to detect bluetongue serotype 8 at an early stage. The heterogeneous sampling scheme in the three countries was taken into account and compared from the angles of given output, herd sensitivity and component sensitivity. For most components, due to the different herd population structure and sampling schemes, herds' sensitivity outputs differed substantially from country to country. However, at component level differences disappeared and countries performed similarly. The output based methods from the comparison between and within countries revealed that in spite of surveillance designs being fairly different, performances seemed to be comparable across countries. In addition, timeliness of detection and the possibility of adding syndromic surveillance were investigated. Syndromic surveillance has increasingly gained interest and it clearly has the advantage of using readily available data for surveillance purposes. However, the study showed that the costs that would be generated by false alarms, as well as indirect costs caused by delays in detection (signals only appear once a significant number of herds are affected), would be higher than any other conventional

surveillance components. Based on these results, one should not rule out the possibility of using these components for emerging or re-emerging diseases but caution is recommended.

Data & Cost for surveillance

Both actions (surveillance and intervention) imply certain costs. For a thorough assessment of surveillance performance and activities, cost of changes must be considered in relation to the added benefit of surveillance (Dehove et al., 2012; Calba et al., 2015; Stärk and Häsler, 2015). Integrating the cost of surveillance into other costs is important to consider too, so one is not blinded by the illusion of conducting surveillance with apparent efficiency, when in fact it is not seen as such when brought to attention at higher levels (Drewe et al., 2014). Though tools developed in chapter III and IV were not explicitly drawn with the aim of reducing cost, they revealed themselves interesting to optimise sample size and design to such an extent that a total cost reduction was allowed while improving surveillance system performance as a whole. In contrast, in chapter VII, the aim was amongst others investigating the best cost efficient alternative for early detection. Sensitivity and specificity of surveillance were used to measure impact on cost. Indeed while sensitivity drives performance, specificity will mainly drive the cost. High performance diagnostic tests could be abandoned for less sensitive diagnostic tests due to lack of specificity (EFSA, 2012). The aim of surveillance could be to spare budgets, but the cost as well as expenditure (stakeholder covering the cost (politician, farmer, and meat industry)) must be considered for sustainable change. Synergies can be found and mutualisation of animal health must be considered; authorities must not be alone in funding. In addition, cost saving should be considered on a larger scale, involving a one health approach to surveillance by taking into account animal welfare, zoonotic potential as well as social and economical impact of disease. Cost efficiency should be evaluated using a cross species and multi component/perspective analysis integrated approach rather than according to a sector specific approach (Häsler et al., 2016; Martins et al., 2016).

In addition, over the last decade, much effort was exercised on improving surveillance systems, while data quality generated by the surveillance system was often considered as an asset. However, the value of information will be hampered by

poor data quality and management. Collecting and analysing surveillance data was a constant challenge whilst compiling this thesis. The lack of consensus regarding motives at a dossier level and/or the interpretation of work results by various laboratories made data exchanges between regional and national laboratories a time consuming task. The lack of standardised formats to report results, to test procedures and diagnostic tests cut off's also lead to difficulties. Thus, as a prerequisite, monitoring data flow, checking data quality, analysing data and ensuring feedback must be carried out in order to provide reliable information. Data generated by surveillance must be recorded and centralised at some level. Means to do so and clear coding and definitions are essential. Also, expected outputs and goals of data records will determine and guide what inputs are needed (FAO, 2011; Stärk and Häsler, 2015).

The challenge while looking into data, not only lies in finding adapted software to analyse terra sized sets of data (i.e., animal movements) or to combine outputs from various databases, or to develop visual tools to view and to represent information gathered from data, it lies mostly in the exercise of inferring from data exploration. The danger of big data comes from overestimating events meaning all assumptions are significant as they diminish variance. Similarly, establishing links between events becomes possible. However, the link does not mean the cause. The rise of molecular typing and the great interest generated by these new tools must be considered with caution. Other factors acting in the triangular interaction (human, host, environment; to which one may add political and social elements) play a crucial role in the spread of diseases. In future, data communication will certainly play a major role and will certainly be a pivotal challenge for surveillance.

REFERENCES

Association Régionale Wallonne de Santé et Identification Animale (ARSIA). Surveillance des avortements. URL (Last update date: 2014): in http://www.arsia.be/Afsca-formations/Rapport-activites/RA-2014-light.pdf (Consulted date: 16/04/2017).

Adkin A, Brouwer A, Simons RRL, Smith RP, Arnold ME, Broughan J, Kosmider R, Downs SH, 2016. Development of risk-based trading farm scoring system to assist with the control of bovine tuberculosis in cattle in England and Wales. Prev Vet Med, 123: 32-38.

Bezos J, Casal C, Romero B, Schroeder B, Hardegger R, Raeber AJ, López L, Rueda P, Domínguez L , 2014. Current ante-mortem techniques for diagnosis of bovine tuberculosis. Res Vet Sci. 97 Suppl:S44-52.

Bisdorff B, Schauer B, Rodriguez-Prieto V, Comin A, Brouwer A , Dorea F, Drewe J, Hoinville L, Lindberg A, Martinez-Aviles M, Martinez-Lopez B, Peyre M, Pinto Ferreira J, Rushton J, Van Schaik G, Stark KD, Staubach C, Vicente-Rubiano M, Witteveen G, Pfeiffer D, Hasler B, 2016. Active animal health surveillance in European Union Member States: gaps and opportunities. Epidemiol Inf , pp. 1–16. doi: 10.1017/S0950268816002697

Bronner A, Hénaux V, Vergne T, Vinard JL, Morignat E, Hendrikx P, Calavas D, Gay E, 2013. Assessing the mandatory bovine abortion notification system in France using unilist capture-recapture approach. PLoS One. 8(5):e63246.

Bronner A, Hénaux V, Fortané N, Hendrikx P, Calavas D, 2014. Why do farmers and veterinarians not report all bovine abortions, as requested by the clinical brucellosis surveillance system in France? BMC Vet Res. 10:93.

Bronner A, Morignat E, Calavas D, 2015. Respective influence of veterinarians and local institutional stakeholders on the event-driven surveillance system for bovine brucellosis in France. BMC Veterinary Research. 11:179. doi:10.1186/s12917-015-0499-1.

Calba C, Goutard FL, Hoinville L, Hendrikx P, Lindberg A, Saegerman C, Peyre M, 2015. Surveillance systems evaluation: a systematic review of the existing approaches. BMC Public Health.15:448.

Calba C, Goutard FL, Vanholme L, Antoine-Moussiaux N, Hendrikx P, Saegerman C, 2016. The Added-Value of Using Participatory Approaches to Assess the Acceptability of Surveillance Systems: The Case of Bovine Tuberculosis in Belgium. PLoS One. 27;11(7):e0159041. doi: 10.1371/journal.pone.0159041. eCollection 2016.

Calvo-Artavia FF, Nielsen LR, Alban L. Epidemiologic and economic evaluation of risk-based meat inspection for bovine cysticercosis in Danish cattle. Prev Vet Med. 2013 Mar 1;108(4):253-61. doi: 10.1016/j.prevetmed.2012.11.002. Epub 2012 Nov 22.

Calvo-Artavia, FF, Alban, L, Nielsen, LR, 2013. Evaluation of surveillance for documentation of freedom from bovine tuberculosis. Agriculture 3, 310–326, http://dx.doi.org/10.3390/agriculture3030310.

Cameron AR, 2012. The consequence of risk-based surveillance: Developing output-based standards for surveillance to demonstrate freedom from disease. Prev Vet Med. 105: 280-286.

Casal C, Alvarez J, Bezos J, Quick H, Díez-Guerrier A, Romero B, Saez JL, Liandris E, Navarro A, Perez A, Domínguez L, Juan Ld, 2015. Effect of the inoculation site of bovine purified protein derivative (PPD) on the skin fold thickness increase in cattle from officially tuberculosis free and tuberculosis-infected herds. Prev Vet Med. 121(1-2):86-92.

Caron A, Grosbois V, Etter E, Gaidet N, de Garine-Wichatitsky M, 2014. Bridge hosts for avian influenza viruses at the wildlife/domestic interface: an eco-epidemiological framework implemented in southern Africa. Prev Vet Med. 117(3-4):590-600. doi: 10.1016/j.prevetmed.2014.09.014. Epub 2014 Oct 18.

Conan A, Goutard FL, Holl D, Ra S, Ponsich A, Tarantola A, Sorn S, Vong S, 2013. Cluster randomised trial of the impact of biosecurity measures on poultry health in backyard flocks. Vet J. 198(3):649-55. doi: 10.1016/j.tvjl.2013.09.010.

Cowie CE, Gortázar C, White PCL, Hutchings MR,Vicente J, 2015. Stakeholder opinions on the practicality of management interventions to control bovine tuberculosis. Vet J 204: 179–185.

Dehove A, Commault J, Petitclerc M, Teissier M, Macé J, 2012. Economic analysis and costing of animal health: a literature review of methods and importance. Rev Sci Tech. 31(2): 605-617.

Drewe JA, Hoinville LJ, Cook AJ, Floyd T, Stärk KD, 2012. Evaluation of animal and public health surveillance systems: a systematic review. Epidemiol Infect. 140(4):575-590.

Drewe JA, Häsler B, Rushton J, Stärk KD, 2014. Assessing the expenditure distribution of animal health surveillance: the case of Great Britain. Vet Rec. 174(1):16.

Drewe JA1, Hoinville LJ, Cook AJ, Floyd T, Gunn G, Stärk KD, 2015. SERVAL: a new framework for the evaluation of animal health surveillance. Transbound Emerg Dis. 62(1):33-45. doi: 10.1111/tbed.12063. Epub 2013 Feb 18.

European Food Safety Authority (EFSA), 2012. Technical meeting of the EFSA Scientific Network for Risk assessment in Animal Health and Welfare - Bovine Tuberculosis Testing. URL (Last update date: 13/12/2012): https://www.efsa.europa.eu/en/supporting/pub/en-288 (Consulted date: 19/02/2017).

European Food Safety Authority (EFSA), 2013. Modelling the impact of a change in meat inspection sensitivity on the surveillance of bovine tuberculosis at country level. URL (Last update date: 28/06/2013): http://www.efsa.europa.eu/en/supporting/pub/450e.htm (Consulted date: 16/10/2015).

European Food Safety Authority (EFSA), 2017. Urgent request on avian influenza opinion. URL (Last update date: 30/01/2017): https://www.efsa.europa.eu/en/efsajournal/pub/4687. (Consulted date: 16:04//2017).

Elbers AR, Gorgievski-Duijvesteijn MJ, Zarafshani K, Koch G, 2010. To report or not to report: a psychosocial investigation aimed at improving early detection of avian influenza outbreaks. Rev Sci Tech. 29(3):435-449.

EUROSTAT, 2013. Milk and dairy production statitics URL (Last update date: 2013): http://ec.europa.eu/eurostat/statistics-explained/index.php/Archive:Milk_and_dairy_production_statistics (Consulted date 20/12/2015).

European Union (EU). Regulation 2016/429 of the European Parliament and of the Council of 2016-03-09 on transmissible animal diseases and amending and repealing certain acts in the area of animal health (Animal Health Law), Official Journal of the European Union L84, 1-208.

FAO, 2011. Challenges of animal health information systems and surveillance for animal diseases and zoonoses. URL (Last update date: 18/12/2011) http://www.fao.org/docrep/014/i2415e/i2415e00.htm (14/07/2014).

Ferrer E, Alfonso P, Ippoliti C, Abeledo M, Calistri P, Blanco P, Conte A, Sánchez B, Fonseca O, Percedo M, Pérez A, Fernández O, Giovannini A, 2014. Development of an active risk-based surveillance strategy for avian influenza in Cuba. Prev Vet Med. 116(1-2):161-167. doi: 10.1016/j.prevetmed.2014.05.012.

Foddai A, Nielsen LR, Willeberg P, Alban L, 2015. Comparison of output-based approaches used to substantiate bovine tuberculosis free status in Danish cattle

herds. Prev Vet Med. 121(1-2):21-9. doi: 10.1016/j.prevetmed.2015.05.005. Epub 2015 May 22.

Goodchild AV, Downs SH, Upton P, Wood JLN, de la Rua-Domenech R, 2015. Specificity of the comparative skin test for bovine tuberculosis in Great Britain. Vet Rec. 177(10): 258. Published online 2015 Sep 2. doi: 10.1136/vr.102961.

Guta S, 2013. Epidemiological investigation of bovine tuberculosis causes of herd breakdowns and persistence in Spain. phD Thesis URL (Last update date: 2013): http://www.tdx.cat/bitstream/handle/10803/125716/sgdb1de1.pdf?sequence=1 (Consulted date: 14/11/2015).

Guta S, Casal J, Napp S, Saez JL, Garcia-Saenz A, Perez de Val B, Romero B, Alvarez J, Allepuz A, 2014. Epidemiological Investigation of Bovine Tuberculosis Herd Breakdowns in Spain 2009/2011. PLoS one 9(8): e104383. doi:10.1371/journal.pone.0104383

Häsler B, Gilbert W, Jones BA, Pfeiffer DU, Rushton J, Otte MJ, 2016. The economic value of One Health in relation to the mitigation of zoonotic disease risks. Curr Top Microbiol Immunol. 2013;365:127-51. doi: 10.1007/82_2012_239.

Hendrikx P, Gay E, Chazel M, Moutou F, Danan C, Richomme C, Boue F, Souillard R, Gauchard F, Dufour B, 2011. OASIS: an assessment tool of epidemiological surveillance systems in animal health and food safety. Epidemiol Infect. 139(10):1486-1496.

Hoinville LJ, Alban L, Drewe JA, Gibbens JC, Gustafson L, Häsler B, Saegerman C, Salman M, Stärk KD, 2013. Proposed terms and concepts for describing and evaluating animal-health surveillance systems. Prev Vet Med. 2013 112(1-2):1-12.

Jajosky RA, Groseclose SL, 2004. Evaluation of reporting timeliness of public health surveillance systems for infectious diseases. BMC Public Health. 2004 Jul 26;4:29.

Jansen J, Wessels R, Lam T, 2012. How to R.E.S.E.T. farmer mindset? Experiences from the Netherlands. Conference paper. Countdown Symposium, Melbourne, Australia.

Knight-Jones TJ, Hauser R, Matthes D, Stärk KD, 2010. Evaluation of effectiveness and efficiency of wild bird surveillance for avian influenza. Vet Res. 41(4):50.

Martin PA, Cameron AR, Greiner M, 2007. Demonstrating freedom from disease using multiple complex data sources 1: a new methodology based on scenario trees. Prev Vet Med. 79(2-4):71-97.

Martins B, Rushton J, Stärk KD, 2016. Economic Assessment of Zoonoses Surveillance in a 'One Health' Context: A Conceptual Framework. Zoonoses Public Health. 2016 Aug;63(5):386-95. doi: 10.1111/zph.12239. Epub 2015 Nov 26.

More SJ, Good M, 2015. Understanding and managing bTB risk: perspectives from Ireland.Vet Microbiol. 176(3-4):209-18. doi: 10.1016/j.vetmic.2015.01.026. Epub 2015 Feb 9.

More SJ, Radunz B, Glanville RJ, 2015. Lessons learned during the successful eradication of bovine tuberculosis from Australia. Vet Rec. 177(9):224-32.

Nielsen PK, Petersen MB, Nielsen LR, Halasa T, Toft N, 2015. Latent class analysis of bulk tank milk PCR and ELISA testing for herd level diagnosis of Mycoplasma bovis. Prev Vet Med. 121(3-4):338-42. doi: 10.1016/j.prevetmed.2015.08.009. Epub 2015 Aug 18.

World Organisation for Animal Health (OIE), 2012. Immediate notification report: Brucellosis, bovine – Belgium B. abortus serovar 3. Report reference: 2012/07621.

Reijn E, Swaan CM, Kretzschmar ME, van Steenbergen JE, 2011.Analysis of timeliness of infectious disease reporting in the Netherlands. BMC Public Health. 11:409. doi: 10.1186/1471-2458-11-409.

Rutten R, Gonzales JL, Elbers ARW, Velthuis AGJ, 2012. Cost Analysis of Various Low Pathogenic Avian Influenza Surveillance Systems in the Dutch Egg Layer Sector. PLoS one 7(4):, http://dx.doi.org/10.1371/journal.pone.0033930.

Schiller I, Oesch B, Vordermeier HM, Palmer MV, Harris BN, Orloski KA, Buddle BM, Thacker TC, Lyashchenko KP, Waters WR, 2010. Bovine tuberculosis: a review of current and emerging diagnostic techniques in view of their relevance for disease control and eradication. Transbound Emerg Dis. 57(4):205-20.

Stärk, KDC, Alonso, S, Dadios, N, Dupuy, C, Ellerbroek, L, Georgiev, M, Hardstaff, J, Huneau-Salaün, A, Laugier, C, Mateus, A, Nigsch, A, Afonso, A, Lindberg, A, 2014. Strengths and weaknesses of meat inspection as a contribution to animal health and welfare surveillance. Food Control 39, 154–162.

Stärk KD, Häsler B, 2015. The value of information: Current challenges in surveillance implementation. Prev Vet Med. 2015 May 14. pii: S0167-5877(15)00188-9. doi: 10.1016/j.prevetmed.2015.05.002. [Epub ahead of print].

van den Berg T, Welby S, Vangeluwe D, Van der Stede Y, Lambrecht B, 2010-2012. Integrating transmission parameters of notifiable Avian Influenza Virus (H5-H7) in Belgium, a surveillance system based on risk (acronym: FLUTREE). N° RF 09/6217 FLUTREE.

Vial F, Thommen S, Held L, 2015. A simulation study on the statistical monitoring of condemnation rates from slaughterhouses for syndromic surveillance: an evaluation based on Swiss data. Epidemiol Infect. 143(16):3423-33. doi: 10.1017/S0950268815000989.

Welby S, Raemaekers M, Vanholme L, Hooyberghs J, Van der Stede Y, 2014. Descriptive analysis of mortality at RENDAC and notified abortions in Belgium during 2012 and 2013. National report from Epidemiology, Risk Assessment and Surveillance Unit of the Veterinairy and Agrochmeical research ve-centre (CODA) to the Federal Agency for the Safety of the Food Chain (FASFC).

Welby, S, Fretin, D, Godfroid, J, Vanholme, L, Hooyberghs, J, Imberechts, H, Van der Stede, Y. Surveillance to declare freedom and/or detect a case? A case study for bovine brucellosis in Belgium. SVEPM Proceedings, SVEPM conference 2015, Ghent, Belgium.

Welby S, Fretin D, Vandewiele H, Imberechts H, Van der Stede Y, 2015. Bovine Tuberculosis Surveillance at Slaughterhouse and Purchase in Belgium: Approach towards Benchmarking. SVEPM Proceedings, SVEPM conference 2015, Ghent, Belgium.

Zurbrigg KJ, Van den Borre NM, 2013. Factors associated with good compliance and long-term sustainability in a practitioner-based livestock disease surveillance system. Can Vet J. 54(3):243-248.

CONCLUSIONS/RECOMMENDATIONS

When evaluating surveillance programs, the key questions are: Is the surveillance effective and is it worth the money?

In this thesis different aspects of the output based methods were explored and its applications to specific issues relating to surveillance were tested, using avian influenza, brucellosis, tuberculosis and bluetongue as case studies. The time that has gone since the changes inspired by these studies were implemented has proven the use and added value of risk based surveillance in disease control programmes and shown how better surveillance could be achieved with less money.

The following conclusions and recommendations can be drawn from these evaluations for the design and assessment of surveillance, notably for official disease surveillance programmes.

Because surveillance will only be effective if those who benefit and/or pay comply with surveillance, a prerequisite prior to surveillance evaluation is for each surveillance programme to have a **clear definition of the objectives and clear communication with** all stakeholders involved, thereby including all respective needs and priorities. Reporting clinical cases is one of the cornerstones of efficient disease surveillance as it will ensure rapid detection and control. Greater incentives, such as compensations, or allowances for veterinarian visits, should be considered; else, setting constraints for non-complying farmers could be a way to increase reporting and increasing surveillance awareness.

Combining different sources of evidence increases the confidence in the country's disease status. Sources of evidence may expand beyond thescope that simple surveillance components encompass, covering that way the whole system from field to fork.

Risk based surveillance is a very attractive alternative, provided knowledge about the main risk factors is available and that one takes into account the fact that they are not necessarily constant over time and across levels (i.e. farm, animal, region, social, political, educational, economical), as well as risk probability of occurrence and consequence. **Data quality** at an individual level, recorded and centralised in a standardised data base, is a cornerstone for effective surveillance. Centralisation of

data on a platform owned by a national organisation should be available and regulated by data sharing protocols in order to respect privacy and ownership issues. **Validation of inputs and outputs** are essential to assess surveillance trends to gauge efficacy and track implemented measures. Imperfect test characteristics should be accounted for in the modelling process as well as impact of design prevalence. Finally, whilst evaluating effectiveness for freedom one can ensure efficacy in case detection simultaneously. Upstream surveillance focussing on probability of risk and downstream surveillance will focus on consequence of risk.

An ideal surveillance system would be one with high **efficacy and low costs.** Although, surveillance activities should be conducted independently and impartially by the public sector to ensure credibility, the public sector alone, cannot support all funding for surveillance activities, especially in countries where the farming industry is increasingly intensive. A minimal financial contribution into a common fund, such as carried out in Belgium by intensive production farms will ensure financial support for surveillance activities on one hand, but, on the other, it will also ensure that compensations can be paid to the sector in case of epidemics.

For high efficacy and low costs, not only sensitivity and direct cost should be accounted for but also specificity and indirect cost of restriction measures applied in surveillance. This surely will impact acceptability and ultimately compliance to surveillance measures. Data quality, representativeness, timeliness during the whole surveillance process whether in active surveillance, syndromic surveillance or passive surveillance remain a challenge and should be assessed in terms of cost efficiency. Finally, expenditure for surveillance should be balanced in relation to other hazards or threats. A one health approach to surveillance, taking into account the disability-adjusted life year (DALY) or quality-adjusted life-year (QALY), should definitely be further investigated to enable a cost efficient approach to surveillance but also societal cost of surveillance in terms of animal welfare.

And to conclude, tools explored during this thesis revealed interesting perspectives for evaluation and design of surveillance taking into account Belgian country specific data and risk factors. Alternatives explored enables conciliating different objectives encountered by the different animal health surveillance stakeholders therby guaranteeing acceptability. Cost could be optimised by improving sensitivity while

taking into account specificity, not only to improve confidence level in freedom but also to enhance disease detection. In addition, consumers and different partners in animal health are getting more and more concerned about the reliability of surveillance effectiveness but also its cost efficiency. Local initiatives are arising for surveillance of non-notifiable diseases, for which mutual trust is desirable to allow for free trade but for sustainable production too. These tools definitely could offer promising perspectives in this context for assessing cost efficiency of other non-regulated diseases and sustainable one health surveillance as a whole across the different levels and sectors encompassed by surveillance.

ACKNOWLEDGEMENTS

I would like to acknowledge everyone who has assisted me throughout my doctoral studies over the years. First, I would like to acknowledge my promotors, Dr Yves van der Stede and Prof Dr Jeroen Dewulf, for agreeing to serve as my promotors and for their patience and feedback during our meetings, skype conversations, email exchanges assistance as I navigated this process to complete my dissertation. Additionally, I would like to thank in particular Dr Koen Mintiens, who initiated me to the field of animal health surveillance at my first start at CODA-CERVA, together with Dr Jozef Hooyberghs from AFSCA, as main sponsor of these studies, for agreeing to serve on my advisory committee. I would like to thank the members of the Jury as well as for their time in reviewing my manuscript and providing valuable feedback at our meeting back in April 2017.

But I also wish to thank the many staff members and colleagues at CODA including Dr Pierre Kerkhofs, Dr Frank Koenen, Dr Herman Van Oyen, Dr Marc Dispas, Xavier Simons, Mickaël Cargnel, Dr Sophie Roelandt, Flavien Riocreux, Geneviève Dekeyser, Dr Valérie Dewaele, Daphné Mercier for their encouragements believing in me during our good and hard moments at CODA. A special thank you for their time in reviewing my Dutch summary is due to Isabelle Behaeghel and Geneviève Dekeyser. Additional gratitude is offered to many colleagues in the field of animal surveillance and risk assessment education who provided resources and feedback during my working carrier at CODA in particular during some of the outbreaks we faced as Tuberculosis, Brucellosis, Bluetongue. This list of individuals includes the following for enlightening me with their immense knowledge: Dr Luc Vanholme, Dr Kris De Clercq, Dr Jacqueline Evers, Dr Thierry Van den Bergh but also of course Dr Jacques Godefroid, Dr Katie Vermeersch, Dr Ellen Stoop, Dr Phillippe Houdart, Dr David Fretin, Dr Virginie Roupie, Dr Jean-François Heymans for taking time out of their busy schedules to review and provide feedback on my studies and scenario tree instrument; Christelle Faes ,Tony Martin for taking the time out of their schedule to provide a statistics tutorial. A special thank is due to my friend and ex-colleague Dr Estelle Meroc for our inspiring lunch walks conversations throughout time at CODA, her support during my thesis preparation and all the many good moments shared at

CODA and beyond. My sincere thanks also go to Sarah Hayes for her support and friendship as well as to my group of schoolmates for all the good times which helped me to distract from my phD.

I also want to thank all of the different persons of the various Belgian animal health stakeholders (ARSIA, DGZ, UPV, VDV, FWA, Boerenbond) who contributed to these studies by providng the data and inputs I needed. I am very grateful to all of them.

I would like to thank and dedicate this thesis to my grandmother, Marie Louise Landuyt Welby. Although it has been years since you have passed, I still take your lessons with me, every day.

Finally, I would like to thank my amazing family for the love, support, and constant encouragement I have gotten over the years. Thank you to my mother Johanna Van der Hoeven for her enthusiasm as I neared my goal and helping me out whenever she could. Thank you to my father Patrick Welby, who originally generated my love for science and who supported me in thoroughly revising my dissertation. But thank you also to my stepfather Stephan Borsen, my mother and father in law Beatrice and Bernard Poncin for their encouragement and continued support over the years. Most importantly, I wish to thank my partner Olivier Poncin for erasing all my doubts and worries, his patience, assistance, supports and faith in me and of course my children Lila and Charly for their good nights, being so brave and help me forget the stress of my phD. Mummy really loves you.

I could not have completed my research without the support of all these wonderful people!

CURRICULUM VITAE

Sarah Welby graduated as a doctor in veterinary medicine in 2004. After an internship at a wildlife veterinary clinic in Namibia and working two years for an animal practice in England, interested in epidemiology and population medicine and undertook a master's in food safety and public health. Following graduation, she immediately was offered a job at the Belgian Veterinary and Agrochemical Research Centre (CODA CERVA), seizing then the opportunities to follow additional courses of the masters in applied statistics. Her daily duties consist in evaluating surveillance systems performance, the impacts of interventions on disease incidence, prevalence and spread, as well as benchmarking surveillance: she thereby provides epidemiological support to the Belgian authorities and animal health stakeholders, for surveillance in Avian Influenza and Salmonella in poultry and Tuberculosis, Brucellosis and Bluetongue in cattle. She had the opportunity to take active part in different European research projects such as EPIZONE, EMIDA, ANIWHA or EFSA working groups which have enabled her to share experience and learn from other countries' approaches in the surveillance of infectious animal disease. She has recently been appointed Head of the Veterinary Epidemiology unit at CODA.

Sarah is author and co-author of various peer reviewed publications in international journals and has presented results of her studies and findings at several national and international congresses.

Written Publications

Welby S. Mechanism of pathogenicity from several bacteria responsible of food transmitted zoonoses, Veterinary University of Liège, Food Microbiology Department, Liège, June 2008, Masters Thesis.

Namata H, **Welby S**, Aerts M, Faes C, Cortiñas Abrahantes J, Imberechts H, Vermeersch K, Hooyberghs J, Mintiens K. Identification of Risk Factors for the Prevalence and Persistence of Salmonella in Belgian Broiler Chicken Flocks, Preventive Veterinary Medicine. 2009 90: 211-222.

Welby S, Namata H, Aerts M, Faes C, Cortiñas Abrahantes J, Imberechts H, Vermeersch K, Hooyberghs J, Mintiens K. Salmonella in Belgian Broiler Chicken Flocks: Identification of Risk Factors for the Prevalence and Persistence. CODA-CERVA Activity Report, 2007-2008.

Welby S, van den Berg T, Marché S, Houdart P, Hooyberghs J, Mintiens K. Evaluation of the belgian active surveillance programme for low pathogenic avian infl uenza in professional poultry holdings. CODA-CERVA Activity Report, 2007-2008.

Welby S, van den Berg T, Marché S, Houdart P, Hooyberghs J, Mintiens K. Redesigning the serological surveillance programme for notifiable Avian Influenza in Belgian professional poultry holdings. Avian Dis. 2010 54(1 Suppl):597-605

Méroc E, **Welby S**, Veldhuis A, Riocreux F, Roelandt S, Vangeel I, Hooyberghs J, Vanholme L, Dispas M, Van der Stede Y. Winterscreening 2009-2010: results for a number of important cattle infectious diseases. CODA-CERVA Activity Report, 2009-10.

Welby S, Poncin O, Claes G, van der stede Y, vangeluwe D, Marché S, Lambrecht B, van den Berg T. Empirical approach for risk based model to enable detection and measures against spread of Low Pathogenic Avian Influenza, preliminary results of wild bird contacts with outdoor poultry pens, CODA-CERVA Activity Report, 2009-2010.

Welby S, Govaerts M; Vanholme L; Mennens K; Hooyberghs J; Maes L; van der Stede Y Bovine tuberculosis in Belgium: Officially free but still sporadic breakdowns! Empirical approach of a risk based surveillance program, CODA-CERVA Activity Report, 2009-2010.

Welby S, Letellier C, Fretin D, Hooyberghs J, Vanholme L, Godefroid J, Van der Stede Y. Evaluation du programme de surveillance pour la Brucellose et la Leucose Bovine enzootique en Belgique. Epidemiol. et Santé Animale, 55, 1-6, Epub 2009.

Welby S, Imberechts H, Riocreux F, Dierick K, Bertrand S, Wildemauwe C, Hooyberghs J, Van der Stede Y. Comparison of Salmonella Enteritidis phage types isolated from layers and humans in Belgium in 2005, Foodborne Pathog Dis. 2011 Aug;8(8):929-34.

Rauw R, Palya V, Van Borm S, **Welby S**, Tatar-Kis T, Gardin Y, Moore Dorsey K, Aly MM, Hassan MK, Soliman MA, Lambrecht B, van den Berg T. Further evidence of antigenic drift and protective efficacy afforded by a recombinant HVT-H5 vaccine against challenge with two antigenically divergent Egyptian clade 2.2.1 HPAI H5N1 strains, Vaccine, 2011, Vaccine. 2011 29(14):2590-600. Epub 2011 Feb 1.

Welby S, Govaerts M, Vanholme L, Mennens K, Hooyberghs J, Maes L, van der Stede Y. Bovine Tuberculosis Surveillance in Belgium: Evaluation of current surveillance components, Épidémiologie et Santé Animale 2011 No. 59/60 pp. 334-336.

Welby S, Govaerts M, Vanholme L, Hooyberghs J, Mennens K, Maes L, Van Der Stede Y. Bovine tuberculosis surveillance alternatives in Belgium. Prev Vet Med. 2012 Sep 15;106(2):152-61.

Vangeel I, Méroc E, Roelandt S, **Welby S,** Riocreux F, Stede YV, Driessche EV, Schoubroeck LV, Czaplicki G, Quinet C, Hooyberghs J, Houdart P, Van der Stede Y. Seroprevalence of Neospora caninum, paratuberculosis and Q fever in cattle in Belgium in 2009-2010. Vet Rec. 2012 Nov 10;171(19):477. doi: 10.1136/vr.101016. Epub 2012 Sep 5.

Vangeel I, Méroc E, Roelandt S, **Welby S**, Riocreux F, Stede YV, Driessche EV, Schoubroeck LV, Czaplicki G, Quinet C, Hooyberghs J, Houdart P, Van der Stede Y. Seroprevalence ofNeospora caninum,paratuberculosis and Q fever in cattle inBelgium in 2009-2010. CODA-CERVA Activity Report, 2011-2012.

Méroc E, **Welby S**, Dewulf J, Depoorter P, Hooyberghs J, Caij AB, Van der Stede Y. Targeted farm sampling in domestic pigs: an adaptation of the belgian aujeszky's disease surveillance program. CODA-CERVA Activity Report, 2011-2012.

Welby S, van der stede Y, Marché S, Lambrecht B, van den Berg T. Risk factors for the introduction of low pathogenic avian influenza virus in Belgian professional poultry holdings: A step towards risk based surveillance. CODA-CERVA Activity Report, 2011-2012.

Welby S, Jasson V, Pierré E, Riocreux F, Vermeersch K, Imberechts H, van der stede Y. Salmonella in Belgian broiler holdings CODA-CERVA Activity Report, 2011-2012.

Baldo A, Leunda A, Do Thi CD, Breyer D, Pauwels K, **Welby S**, Van Vaerenbergh B, Herman P. Biosafety Risk Assessment and Management of Laboratory-derived Influenza A (H5N1) Viruses Transmissible in Ferrets. Applied Biosafety 2013 18:6-17.

Claes G, **Welby S**, Van Den Berg T, Van Der Stede Y, Dewulf J, Lambrecht B, Marché S. The impact of viral tropism and housing conditions on the transmission of three H5/H7 low pathogenic avian influenza viruses in chickens. Epidemiol Infect. 2013 141(11):2428-43. doi: 10.1017/S0950268813000125. Epub 2013 Feb 11.

Welby S, Méroc E, Faes C, De Clercq K, Hooyberghs J, Mintiens K, Van der Stede Y. Bluetongue surveillance system in Belgium: a stochastic evaluation of its risk-based approach effectiveness. Prev Vet Med. 2013 Oct 1;112(1-2):48-57. doi: 10.1016/j.prevetmed.2013.07.005. Epub 2013 Aug 12.

Ensoy C, Aerts M, **Welby S**, Van der Stede Y, Faes C. A Dynamic Spatio-Temporal Model to Investigate the Effect of Cattle Movements on the Spread of Bluetongue

BTV-8 in Belgium. PLoS One. 2013 Nov 11;8(11):e78591. doi: 10.1371/journal.pone.0078591.

Ensoy C, Faes C, **Welby S**, Van der Stede Y, Aerts M.-Exploring Cattle Movements in Belgium. Article in Preventive Veterinary Medicine 116(1-2) · September 2014.

Rauw F, Van Borm S, **Welby S**, Ngabirano E, Gardin Y, Palya V, Lambrecht B. Quantification of rHVT-F genome load in feather follicles by specific real-time qPCR as an indicator of NDV-specific humoral immunity induced by day-old vaccination in SPF chickens Avian Pathology 2015 44(3):1-31.

Verpoest S, Brigitte Cay A, Van Campe W, Mostin L, **Welby S**, Favoreel H, De Regge N. Age and strain dependent differences in the outcome of experimental infections of domestic pigs with wild boar pseudorabies virus isolates. J Gen Virol 2015 Nov 20. Epub 2015 Nov 20.

Welby S, van Schaik G, Veldhuis A, Brouwer H, Peroz C, Berends I, Fourichon C, Wever P, van der Stede Y.Cost efficiency of passive active and syndromic surveillance components fore earlydetection of emerging vector borne disease in cattle: Bluetongue as case study for Belgium, France and the Netherlands. CODA-CERVA Activity Report, 2013-2014.

Welby S, Méroc E, Faes C, De Clercq K, Hooyberghs J, Mintiens K, Van der Stede Y. Bluetongue surveillance system in Belgium: a stochastic evaluation of its risk-based approach effectiveness. CODA-CERVA Activity Report, 2013-2014.

Veldhuis A, Brouwer-Middelesch H, Marceau A, Madouasse A, Van der Stede Y, Fourichon C, **Welby S**, Wever P, van Schaik G. Application of syndromic surveillance on routinely collected cattle reproduction and milk production data for the early detection of outbreaks of Bluetongue and Schmallenberg viruses. Prev Vet Med. 2016 Feb 1;124:15-24. doi: 10.1016/j.prevetmed.2015.12.006. Epub 2015 Dec 18.

Michiels T, **Welby S**, Vanrobaeys M, Quinet C, Rouffaer L, Lens L, Martel A, Butaye P. Prevalence of Mycoplasma gallisepticum and Mycoplasma synoviae in commercial poultry, racing pigeons and wild birds in Belgium. Avian Pathol. 2016 Jan 26:1-28. [Epub ahead of print].

Welby S, van Schaik G, Veldhuis A, Brouwer H, Peroz C, Berends I, Fourichon C, Wever P, van der Stede Y.Cost efficiency of passive active and syndromic surveillance components fore earlydetection of emerging vector borne disease in cattle: Bluetongue as case study for Belgium, France and the Netherlands. Transbound Emerg Dis. 2016 Sep 26. doi: 10.1111/tbed.12564. [Epub ahead of print].

Oral presentations:

Welby S, Mechanism of pathogenicity from several bacteria responsible of food transmitted zoonoses, Veterinary University of Liège, Food Microbiology Departement, Masters Thesis oral defense, Liège, Belgium, June 2008.

Welby S, van den Berg T, Marché S, Houdart P, Hooyberghs J, Mintiens K. Evaluation of Avian Influenza active surveillance program in professional poultry holdings in Belgium, Journée Communication AFSCA-CERVA, Brussels, Belgium, May 2008.

Welby S, van den Berg T, Marché S, Houdart P, Hooyberghs J, Mintiens K. Redesigning the serological surveillance program for notifiable Avian Influenza in Belgian professional poultry holdings, EPIZONE, Brescia, Italy, June 2008.

Welby S, van den Berg T, Marché S, Houdart P, Hooyberghs J, Mintiens K. Evaluation of Avian Influenza active surveillance program in professional poultry holdings in Belgium, VEE, Leuven, Belgium, October 2008.

Welby S, van den Berg T, Marché S, Houdart P, Hooyberghs J, Mintiens K. Redesigning the serological surveillance program for notifiable Avian Influenza in Belgian professional poultry holdings, 7th International Symposium on Avian Influenza, Athens, Florida, 2009.

Welby S, Namata H, Aerts M, Faes C, Cortiñas Abrahantes J, Imberechts H, Vermeersch K, Hooyberghs J, Mintiens K, Salmonella in Belgian broiler chicken flocks: Identification of risk factors for the prevalence and persistence, Journée Communication AFSCA-CERVA, Brussels, Belgium, May 2009.

Welby S, Méroc E, Faes C, Mintiens K, De Clercq K, Hooyberghs J, Van der Stede Y, EPIZONE: Epidemiology and surveillance of BTV-8 in Europe : Evaluation of BTV-8 surveillance in Belgium, EPIZONE, Antalya, Turquie, May 2009.

Welby S, Letellier C, Fretin D, Hooyberghs J, Vanholme L, Godefroid J, Van der Stede Y. Evaluation du programme de surveillance pour la Brucellose et la Leucose Bovine enzootique en Belgique, AESA-AEMA meeting, Paris, France, June 2009.

Welby S, Mintiens K, van den Berg T, Fretin D, Marché S, Letellier C, Houdart P, VanHolme L, Hooyberghs J, Godefroid J, van der Stede Y. Avian Influenza, Bovine Brucellosis and Bovine Leucosis in Belgium: Evaluation of the Most Optimal Sampling Strategy through Complex Surveillance Data Analysis, ISVEE, Durban, South Africa, August 2009.

Welby S, Méroc E, Faes C, Mintiens K, De Clercq K, Hooyberghs J, Van der Stede Y. Risk approach models to estimate the sensitivity for different surveillance

components: Bluetongue in Belgium a case study, VEE, Tervuren, Belgium, October 2009.

Welby S, Course for Bio Ingénieurs : Epidémiosurveillance en santé animale, Brussels, Belgium, December, 2009.

Welby S, Course for qualified vets Masters Veterinary Public Health: Scenario tree modeling, Liège, Belgium, November, 2010.

Claes G, **Welby S**, Lambrecht B, Marché S, Dewulf J, Van der Stede Y, Van den Berg T. Low Pathogenic H5 and H7 Avian Influenza: Assessment of Transmission Parameters and the Influence of Housing Conditions on Transmission, VEE, Roselaere, October 2010.

Welby S, Govaerts M, Riocreux F, Vanholme L, Hooyberghs J, Mennens K, Maes L, Van Der Stede Y. Bovine Tuberculosis in Belgium: Officially free but …?! Empirical approach for a risk based surveillance program, SVEPM, March 2011.

Welby S, Govaerts M, Riocreux F, Vanholme L, Hooyberghs J, Mennens K, Maes L, Van Der Stede Y. Bovine Tuberculosis in Belgium: Empirical approach for a risk based surveillance program, EPIZONE, Arnhem, April 2011.

Welby S, Govaerts M, Riocreux F, Vanholme L, Hooyberghs J, Mennens K, Maes L, Van Der Stede Y. Bovine tuberculosis in Belgium: Empirical approach for a risk based surveillance program, VEE, St Niklaas, October 2011.

Claes G, Marché S, **Welby S**, Lambrecht B, Dewulf J, Van der Stede Y, Van den Berg T, 2011. Low Pathogenic H5 and H7 Avian Influenza: Reproduction Ratio and Influence of Housing Conditions on Transmission, VEE, St Niklaas, October 2011.

Welby S, Marché S, Poncin O, Claes G, Vangeluwe D, Lambrecht B, van den Berg T, Van der Stede Y, 2012. Impact of wild waterfowl incursions and holding characteristics: Astep towards Avian Influenza virus risk based surveillance in Belgian. ISVEE, Maastricht August 2012.

Welby S, Fretin D, Vandewiele H, Imberechts H, Van der Stede Y, 2015. Bovine Tuberculosis Surveillance at Slaughterhouse and Purchase in Belgium: Approach towards Benchmarking. VEE Melle, October 2015.

Welby S, De Clercq K, 2016. Effectiveness and cost efficiency of surveillance for proving freedom & early detection of an emerging disease: Case study comparing different surveillance schemes in different countries, FAO, EU FMD Webinars 20/01/2016.

Poster Presentations:

Welby S, van den Berg T, Marché S, Houdart P, Hooyberghs J, Mintiens K, Redesigning the serological surveillance programme for notifiable Avian Influenza in Belgian professional poultry holdings, SVEPM, London, March 2009.

Welby S, Letellier C, Fretin D, Hooyberghs J, Vanholme L, Godefroid J, Van der Stede Y. The use of Scenario Tree methodology to evaluate the different sampling components in the Belgian Surveillance programme for Brucellosis, VEE Brussels, Belgium, October 2009.

Welby S, Méroc E, Faes C, Mintiens K, De Clercq K, Hooyberghs J, Van der Stede Y. Bluetongue Surveillance in Belgium in 2007: Evaluation of the different surveillance components, AESA, Liège, Belgium, May 2010.

Welby S, Méroc E, Faes C, Mintiens K, De Clercq K, Hooyberghs J, Van der Stede Y. Bluetongue Surveillance in Belgium in 2007: Evaluation of the different surveillance components, EPIZONE, St Malo, France, June 2010.

Welby S, Poncin O, Claes G, Van der Stede Y, Vangeluwe D, Marché S, Lambrecht B, van den Berg T . Flutree: An empirical approach for risk based model to enable detection and measures against spread of Low Pathogenic Avian Influenza, VEE, Roselaere, October 2010.

Welby S, Govaerts M, Riocreux F, Vanholme L, Mennens K, Hooyberghs J, Maes L, van der Stede Y. Bovine Tuberculosis in Belgium: free or not free? Evaluation of current surveillance components, ICAHS, Lyon, 2011.

Welby S, Poncin O, Claes G, Van der Stede Y, Vangeluwe D, Marché S, Lambrecht B, van den Berg T. Flutree: An empirical approach for risk based model to enable detection and measures against spread of Low Pathogenic Avian Influenza, BWDS, CODA Tervuren, October 2011.

Welby S, van den Berg T, Lambrecht B, van der Stede Y, Marché S. Avian Influenza virus in Belgian poultry: Identifying risk farms, a step towards risk based surveillance, International 8th International Symposium on Avian Influenza Avian Influenza in Poultry and Wild Birds 1 – 4 April 2012Royal Holloway, University of London, UK.

Welby S, Fretin D, Vandewiele H, Imberechts H, Van der Stede Y, 2015. Bovine Tuberculosis Surveillance at Slaughterhouse and Purchase in Belgium: Approach towards Benchmarking. SVEPM Proceedings, SVEPM conference 2015, Ghent, Belgium.

Welby, S, Fretin, D, Godfroid, J, Vanholme, L, Hooyberghs, J, Imberechts, H, Van der Stede, Y. Surveillance to declare freedom and/or detect a case? a case study for

bovine brucellosis in Belgium. SVEPM Proceedings, SVEPM conference 2015, Ghent, Belgium.