



Risk based surveillance for vector-borne diseases in horses : combining multiple sources of evidence to improve decision making

Céline Faverjon

► To cite this version:

Céline Faverjon. Risk based surveillance for vector-borne diseases in horses : combining multiple sources of evidence to improve decision making. Agricultural sciences. Université Blaise Pascal - Clermont-Ferrand II, 2015. English. NNT : 2015CLF22604 . tel-01333559

HAL Id: tel-01333559

<https://theses.hal.science/tel-01333559>

Submitted on 17 Jun 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Blaise Pascal

Université d'Auvergne

N°D.U. 2604

Année 2015

Ecole Doctorale Sciences De La Vie, Santé, Agronomie, Environnement

N° d'ordre : 674

THÈSE

Présentée à VetAgro Sup, Lyon

Pour l'obtention du grade de

DOCTEUR D'UNIVERSITÉ

(Spécialité : Épidémiologie)

Soutenue le 15 octobre 2015

par

Céline FAVERJON

Risk based surveillance for vector-borne diseases in horses: combining multiple sources of evidence to improve decision making

Jury

Mme STAERK Katharina
Mr BAYLIS Matthew
Mr BIRON David
Mr FISCHER Egil
Mme LEBLOND Agnès
Mme DE KOEIJER Aline
Mr HENDRIKX Pascal

Royal Veterinary College, Londres
Université de Liverpool
CNRS, Université Blaise Pascal
Université d'Utrecht
INRA, VetAgro Sup, Lyon
Central Veterinary Institut, Lelystad
ANSES

Rapporteur
Rapporteur
Président
Membre invité
Directrice de thèse
Co-directrice de thèse
Encadrant

INRA - Unité d'Épidémiologie Animale UR 0346

*À Marcel et Roger,
qui n'auront pas vu la fin de ce travail.*

ACKNOWLEDGEMENTS

First of all, I would like to thank the members of the jury, *Katharina Stärk*, *Matthew Baylis* and *David Biron*, for having accepted the invitation to the defense of my doctoral thesis and for reading and commenting this document.

Of course I also would like to express my thanks to my supervisors who offered me the opportunity of this PhD and who guided me throughout this work. Thank you to *Aline de Koeijer* for her advices and valuable comments especially regarding the tricky task of risk assessment. I am especially thankful for her frankness and her warm welcome in the Netherlands. Merci à *Agnès Leblond* pour m'avoir toujours accompagnée et encouragée tout au long de ce travail. Merci aussi pour la confiance qu'elle m'a accordée et la liberté de choix qu'elle m'a laissée durant ces 3 années.

I also would like to thank all the members of the EMIDA-VICE project for the constructive and valuable interactions throughout this work. They contributed to my reflection as a Ph.D. student, but I am sure that they will also influence my future work. Particular thanks must be recorded to *Egil Fischer* for his support, advice and availability during these 3 years. Thank you also for our many discussions from statistics to Dutch and French culture. I also want to express my thanks to *Gunnar Andersson* for his great help with the Bayesian approach and statistics, for his advice and his kindness. Thank you also to *René Bødker* for his kind welcome in Denmark. Finally, special thanks to *Flavie Vial* for her help during this work especially during my “discovery” of the syndromic surveillance world.

In addition, I would like to thank all the institutes who provided the data used in this work and all persons involved in data collection networks. Without this essential contribution, the majority of activities outlined in this work would never have been realized. Particular thanks to the *Ministère de l'agriculture français* and the *Direction Générale de l'Alimentation* for having provided data on mortality in horses and on trade in living animals (a thank you to *Pascal Hendriks* who was a perfect facilitator), to *RESPE* for the supply of data on nervous and respiratory symptoms in horses, to *SAGIR* for having provided data on mortality in wild birds, to *MARS-Agri4cast* at the *Institute for Environment and Sustainability*, to the *European Commission* for the supply of temperature data, and to *CIRAD* with the *EID-Med*, and the *IPPTS* with the support of the *Directions départementales de la protection des population*, who collected and shared entomological data.

This research was carried out as part of the VICE project funded under the FP7 EMIDA ERA-NET initiative. National funding within ERA-Net was provided by the *Dutch Ministry of Economic Affairs* (Project nr. BO-20-009-009). This work was also partially financially supported by *IFCE* and *VetAgro Sup*. I would like to thank all these institutes for having made this work possible.

Je tiens à remercier mes « collègues de couloir » pour nos nombreuses pauses café/gâteaux partagées. Merci notamment à *Véronique* la maman attentive, et à *Luc* pour le bureau de ministre. Un merci tout particulier à *Alice* pour sa sympathie et nos discussions multiples. Ces années n'auraient très certainement pas été pareil sans toi dans ce bout de couloir. Je remercie également mes « collègues de Kfet », *Bernadette* et *Caroline*, pour les litres de café-philo avalés durant ces 3 années.

Un grand merci à la petite troupe des doctorants de VetAgro Sup pour nos espoirs et désespoirs en tout genre. Merci plus particulièrement à *Julie* ma chère co-bureau, à *Sylvain* mon co-représentant, et à *Anne-Laure* ma co-sushi-ciné.

Je remercie également l'Unité EpiA pour les accueils sympathiques lors de mes passages clermontois et pour l'aide fournie lorsque j'en ai eu besoin. Merci en particulier aux (ex-)doctorants pour nos échanges et à *Françoise* pour sa disponibilité lors de mes OM multiples.

I would like to thank the team at CVI for the nice four months spent together "in the middle of nowhere". Special thanks to *Jantien* for her kindness and her smile (especially when I tested the "weird" Dutch cuisine!).

Pour finir, je voudrais remercier ma famille pour son indéfectible soutien et mes amis pour ces belles années passées à leurs côtés (il y en aura d'autres, je n'en doute pas !). Mes remerciements tout particuliers à *Laure* pour son œil aiguisé et attentif tout au long de ce travail (et pour tout le reste bien sûr aussi, surtout !).

Ich danke *Malte* für seine Geduld, seine Unterstützung und für un peu trop de choses pour pouvoir être exprimées ici et en allemand.

CHAPTER I: INTRODUCTION 1

A. CONTEXT AND NEEDS.....	1
1. Vector-borne diseases, a major concern	1
2. Equine industry, risk and specificities.....	1
2.1. Population at risk.....	1
2.2. Population not well tracked	2
3. Early warning.....	2
3.1. Active surveillance	3
3.2. Passive surveillance	3
3.3. Early warning in horse population.....	3
4. Risk assessments	4
5. Syndromic surveillance.....	5
B. RESEARCH QUESTION	7
C. CASES STUDIES	9
1. French equine industry.....	9
1.1. Equine population	9
1.2. Organization of the equine industry.....	9
1.3. Disease surveillance in French horses	10
2. Diseases of interest	11
2.1. African horse sickness	11
2.2. Equine encephalosis	12
2.3. West Nile virus.....	12
D. OUTLINES OF THE WORK	15

CHAPTER II: QUANTITATIVE RISK ASSESSMENTS 17

A. OVERVIEW.....	19
1. General principle of risk assessment	19
1.1. Definitions and objectives	19
1.2. Method	20
2. Probability of entry.....	21
2.1. Definition	21
2.2. Estimation for VBDs.....	21
3. Probabilities of transmission and establishment	22
3.1. Probability of transmission	22
3.2. Probability of establishment.....	23
4. Conclusion	24
B. PROBABILITIES OF ENTRY AND ESTABLISHMENT	25
1. Introduction.....	25
2. Paper 1.....	26
C. COMPARISON OF DISEASES RISKS	41
1. Introduction.....	41
2. Method	42

2.1.	Model for risk assessment of viruses entry	42
2.2.	Parameters	42
2.3.	Input data	43
2.4.	What-if scenarios.....	44
2.5.	Calculations.....	44
3.	Results	45
3.1.	Spatiotemporal probability of entry	45
3.2.	Sensitivity analysis	47
3.3.	What-if scenarios.....	48
4.	Discussion	49
D.	DISCUSSION AND CONCLUSION	51
1.	Discussion	51
2.	Conclusion	52

CHAPTER III: SYNDROMIC SURVEILLANCE..... 53

A.	OVERVIEW	55
1.	Overall principle of syndromic surveillance.....	55
1.1.	History.....	55
1.2.	Definition and objectives	56
1.3.	Overall approach	56
2.	Preliminary steps	57
2.1.	Definition of objectives.....	57
2.2.	Overview of data available	57
2.3.	Definition of syndrome	58
3.	Data description and preprocessing	58
3.1.	Data description.....	58
3.2.	Data preprocessing	58
4.	Detection algorithms	59
4.1.	Choice of detection algorithm	59
4.2.	Historical limits	61
4.3.	Control charts	61
4.4.	Regression methods	62
5.	Assessment of performance	64
5.1.	Performance metrics	64
5.2.	Test data	65
6.	Conclusion	66
B.	COMPARISON OF PRE-PROCESSING METHODS	67
1.	Introduction	67
2.	Methods.....	68
2.1.	Data characterization.....	68
2.2.	Data pre-processing.....	68
2.3.	Forecasting	70
2.4.	Detection algorithm.....	70
2.5.	Quantitative assessment	72
2.6.	Implementation	72
3.	Results	72

3.1.	Baseline characterization	72
3.2.	Smoothing and forecasting	73
3.3.	Outbreak detection	76
4.	Discussion	79
5.	Conclusion	81
C.	VALUE OF EVIDENCE	83
1.	Introduction	83
2.	Paper 2	84
D.	DISCUSSION AND CONCLUSION	93
1.	Discussion	93
2.	Conclusion	94
CHAPTER IV:	MULTIPLE INDICATORS OF RISK	95
A.	MULTISTREAM SYNDROMIC SURVEILLANCE	97
1.	Introduction	97
2.	Paper 3	98
B.	COMBINING QUANTITATIVE RISK ASSESSMENT AND SYNDROMIC SURVEILLANCE	111
1.	Introduction	111
2.	Material and Methods	112
2.1.	Bayesian framework	112
2.2.	Simulated EE outbreaks	113
2.3.	Probability of EE introduction	115
2.4.	Syndromic surveillance of EE	115
2.5.	Presentation of the concept and first feedback	116
3.	Results	116
3.1.	Multivariate syndromic surveillance	116
3.2.	Combining Risk assessment	117
3.3.	Workshop output	117
4.	Discussion and Conclusion	119
C.	DISCUSSION AND CONCLUSION	121
1.	Discussion	121
2.	Conclusion	121
CHAPTER V:	DISCUSSION	123
1.	Early warning system of VBDs in horses	123
1.1.	Quantitative risk assessment	123
1.2.	Syndromic surveillance	124
1.3.	Combining risk	124
2.	Support decision-making	125
3.	Demonstrate freedom of disease	126
4.	Practical implementation of integrated surveillance systems	127
5.	Conclusion	128
APPENDICES		131
REFERENCES		179

List of Figures :

Figure 1: Estimated geographical distribution of horses population (ICFE-SIRE)	9
Figure 2: Geographic distribution of the French sentinel veterinarians involved in RESPE.	11
Figure 3 : National median probability of EE and AHS viruses entry to France.....	45
Figure 4 : Median probabilities of viruses entry via an infectious host. Example of the year 2012.	46
Figure 5 : Median probabilities of viruses entry via an infectious vector. Example of the year 2012.....	46
Figure 6 : At risk areas for viruses entry when entry via an infectious host and via an infectious vector are considered together.	47
Figure 7 : Correlation of the model input parameters with the probability of entry of AHS and EE via an infectious host	47
Figure 8 : Correlation of the model input parameters with the probability of entry of AHS and EE via an infectious vector.....	48
Figure 9 : Overall method to implement syndromic surveillance.	57
Figure 10 : Number of nervous symptoms cases declared to RESPE from 2006 to 2013.....	68
Figure 11 : Four time series used	69
Figure 12 : West Nile virus outbreaks.....	71
Figure 13 : Two examples of simulated outbreaks inserted in TS0. Positions of outbreaks identified with dotted lines above the peak.	71
Figure 14 : Decomposition of time series.....	73
Figure 15: Combining syndromic surveillance and probability of disease introduction to obtain a posterior probability that there is an ongoing outbreak.	111
Figure 16: EE simulated outbreak in Aquitaine	114
Figure 17: Value of evidence in France at week 31 in year 2012 from the simulated outbreak in Aquitaine, using a detection radius of 50 km	116
Figure 18: Posterior probability of EE outbreak in France in year 2012 during the simulated outbreak in Aquitaine from week 29 to week 32	118

List of Tables :

Table 1: Main players in the French equine industry	10
Table 2: Mandatory notifiable diseases to the French ministry..	10
Table 3 : import procedure for African horse sickness and equine encephalosis	42
Table 4 : Parameters specific of diseases and used for African horse sickness and Equine encephalosis in the model of viruses entry.....	43
Table 5 : Description of scenario tested. Scenarios 1 to 5 changed the existing legislation, scenarios 6 to 7 changes imports from third countries.	44
Table 6 : Change (in %) on median probabilities of viruses entry to France compared to the default scenario.	49
Table 7 : Summary of the advantages and limitations of the main detection algorithms in the time series analysis.	60
Table 8 : Metrics used for evaluation of outbreak detection algorithms	65
Table 9 : Advantages and disadvantages of types of test data (Buckeridge et al. 2005)	66
Table 10 : Smoothing and forecasting performance of GLMs..	74
Table 11 : Smoothing and forecasting performance of Holt-Winters models.....	75
Table 12 : System's detection performances	77
Table 13 : system's detection performances for time series fitted with generalized linear models.....	78
Table 14: system's detection performances for time series fitted with Holt-Winters	79

List of Appendices:

Appendix 1: Model calculation for PW-host. Details of calculation regarding the AHSV introduction via the import of an infectious host.	133
Appendix 2: Model calculation for PW-vector. Details of calculation regarding the AHSV introduction via the import of an infectious vector.....	141
Appendix 3: Model parameters. Description of all parameters used in the model calculation for PW-host and PW-vector.....	143
Appendix 4: Article in a Professional journal published in 'Bulletin épidémiologique du RESPE'. Juillet 2013. C. Faverjon, S. Lecollinet, S. Zientara, A. Leblond. 'Peste équine, quel risque pour la France ?' (http://www.respe.net/node/1604).	145
Appendix 5: Article in a Professional journal published in 'N° spécial de Pratique Vétérinaire Equine'. Octobre 2014, 142-151. S. Zientara, C. Faverjon, A. Leblond, S. Lecollinet. La peste équine : épidémiologie, diagnostic et prévention.....	147
Appendix 6: Poster presented at the 7th EPIZONE meeting in Brussel, 2013.	157
Appendix 7: Poster presentation at the Conference of the Society for Veterinary Epidemiology and Preventive Medicine (SVEPM) in Ghent, 2015.	159
Appendix 8: Poster presentation at the Journées de l'Ecole doctorale SVSAE in Clermont-Ferrand, 2015.	161
Appendix 9: Poster presentation at the Conference of the Society for Veterinary Epidemiology and Preventive Medicine (SVEPM) in Ghent, 2015.	163
Appendix 10: Article in a Professional journal published in 'Bulletin Epidémiologie et Santé Animale' by AEEMA. 2015. Épidémiol. et santé anim., 2015, 67, 87-95	165
Appendix 11: EEV simulated outbreak in Normandie.	175
Appendix 12: Posterior probability of EE outbreak in France in year 2012 during the simulated outbreak in Normandie from week 36 to week 41. Radius for detection is 50 km (Andersson et al. 2015).....	177

List of abbreviations

ACF: autocorrelation function

AIC : Akaike information criterion

AHS : African horse sickness

ARIMA : autoregressive integrated moving average

ARMA : Autoregressive moving average models

AUC : Area Under the receiver operating characteristic curve

BTV : Bluetongue virus

CUSUM : cumulative sums

EE : Equine encephalosis

EHV : Equine herpes virus

EU : European Union

EWMA : exponential weighted moving average

FAO : Food and Agriculture Organization of the United Nations

GLMs: generalized linear regression models

HW: Holt-Winters generalized exponential smoothing

IFCE : Institut français du cheval et de l'équitation

LMs : Linear models

NB : Negative Binomial

OIE : World organization for Animal health

PACF : partial autocorrelation function

R_0 : basic reproductive number

RESPE : French network for the surveillance of equine diseases

RMSE : root-mean-squared error

SARIMA : seasonal ARIMA

SIRE : “système d’information relatif aux équidés »

TRACES : Trade Control and Expert System

TS : time series

USA : United States of America

VBDs : vector-borne diseases

WHO: World health organization

WNV : West Nile virus

CHAPTER I: INTRODUCTION

A. CONTEXT AND NEEDS

1. Vector-borne diseases, a major concern

The incidence of emerging infectious diseases has increased during recent decades due to changing socio-economic, environmental, and ecological factors (Jones et al. 2008). This is especially true for vector-borne diseases (VBDs), which represent almost 29% of emerging infectious diseases in the last decade (Jones et al. 2008). The introduction and spread into Europe in the 1979 of *Aedes albopictus*, an efficient vector for at least 22 arboviruses (Gratz 2004), of bluetongue virus (specifically BTV-8) in 2006 (Zientara and Sánchez-Vizcaíno 2013), and Schmallenberg virus in 2011 (Doceul et al. 2013) are good examples of the problem posed by VBDs. The specific emergence of VBDs is probably due in part to climate anomalies that occurred during the 1990s (Jones et al. 2008), as has been shown for BTV in Europe (Guis et al. 2012; Purse et al. 2005). Indeed, as explained by Massad and colleagues (Massad et al. 2011), “increased spread of insect-borne diseases is likely in a warmer world”. In the context of global warming, emerging VBDs are thus a growing concern, all the more so because of their huge economic and social impact (Marsh et al. 2008). To address this challenge, Lindgren and colleagues (Lindgren et al. 2012) argue in favor of developing novel approaches for risk assessment and surveillance in order to enhance preparedness and to facilitate public-health decision making.

2. Equine industry, risk and specificities

2.1. Population at risk

Horses, unlike livestock, typically travel frequently over short and long distances around the world for competition, training, and/or reproduction. These movements increase the risk of the dissemination of infectious diseases (Robin et al. 2011). This is a concern not just for the equine industry but also for public health. Indeed, numerous equine viruses are zoonotic (e.g., rabies, brucellosis, anthrax, glanders, leptospirosis, Hendra virus). However, controlling diseases spread by equines is not only important from a sanitary point of view but also with regard to the important economic weight of the equine industry, particularly in Europe (Liljenstolpe 2009). As an example, in 2010, the European equine industry encompassed 3.7 million horses, generated 100 billion euros a year, and provided the equivalent of 400,000 full-time jobs (Leadon and Herholz 2009). Furthermore, the sector is growing, with an increase in the number of horse riders of 5% per annum. The introduction of exotic infectious disease may thus have

huge economic consequences, as was seen with the 13-week outbreak of African horse sickness in Portugal in 1990, whose total cost was estimated around US \$2 million (Portas et al. 1999).

2.2. Population not well tracked

Despite the sanitary and economic impacts of equine diseases, effective health regulations and biosecurity systems to ensure safe equine movements are not always in place at the national and international level (Leadon and Herholz 2009; Murray et al. 2013). This was illustrated by the outbreak of equine influenza in Australia in 2007. Here, the authorities failed to contain the infection in quarantine following the importation of one or more infected horses (Webster 2011). The horse population is also not well-tracked, which complicates the control and surveillance of diseases. In the EU, the implementation of mandatory passports for horses in 2008 has improved the tracking of horses. However, the database that contains the information on animal movements and deaths is not regularly updated. This has two consequences. Firstly, the exact number of horses and their geographical location is unknown, which is an obstacle for disease surveillance and control. Secondly, the exact number of horses transported between EU member states or within a country is still not available. Indeed, although the EU's Trade Control and Expert System (TRACES) (Commission Decision 2003) provides information on the number of horses imported to and within the EU, several movements are not recorded in the database due to the absence of mandatory transport notification.

The difficulties of implementing proper health regulations and tracking systems can be explained by the complex reality of the world of horses. The equine industry includes a myriad of activities (e.g., tourism, equestrian sports, breeding and slaughtering of horses) and the various stakeholders engaged in each activity do so with different expectations, ranging from professional to leisure (Castejón-Montijano and Rodríguez-Fernández 2011).

Considering the potential health and financial risks posed by horses, it is especially important to develop novel approaches for the surveillance of exotic infectious diseases, such as VBDs. However, this also constitutes an additional challenge given the structure of the equine industry.

3. Early warning

A key point in controlling emerging or reemerging VBDs is early warning. Indeed, dealing with a disease outbreak in its early stages is easier and more economical than once it has become widespread (FAO-OIE-WHO Collaboration 2013; FAO-OIE-WHO Collaboration 2006). Early warning systems are timely surveillance systems aimed at predicting the probability that an outbreak is spreading to new areas in order to trigger prompt public health interventions (FAO-OIE-WHO Collaboration 2006). Different strategies such as active and/or passive surveillance are used to ensure the timeliness of detection.

3.1. Active surveillance

Active surveillance refers to the active role of health authorities in data collection. The advantage is that active sampling may detect a disease without the observation of clinical signs. However, one of the major drawbacks is that, to detect rare diseases like a newly introduced exotic disease, active sampling has to be very large and redundant, which can be very costly (Doherr and Audigé 2001). To mitigate costs, it is possible to implement a specific type of active surveillance, known as risk-based surveillance. Risk-based surveillance is defined by Stärk and colleagues (Stärk et al. 2006) as the allocation of surveillance activities based on the probability of events with or without consideration of the consequences of the event, the management of the event, or the process of communication of the event. The term “targeted surveillance” is also used (Doherr and Audigé 2001; de Koeijer et al. 2002; Salman 2003). The simple idea behind the concept is to collect samples from the segments of the population that have the highest probabilities of being infected, thus increasing the probability of disease detection (Salman 2003). In this way, this process decreases the required sample size without reducing the probability of detecting the disease. Risk-based surveillance is based on the concept of looking for something where it is most likely to be found; this approach thus requires reliable and thorough prior information on at-risk populations in order to ensure the appropriate representativeness of the sampling (FAO 2014; Oidtmann et al. 2013; Stärk et al. 2006).

3.2. Passive surveillance

In many countries, passive surveillance is one of the most common forms of surveillance for rare and exotic diseases. The term refers to any passive disease reporting systems in which veterinarians, farmers, or any other stakeholders notify authorities when they have sick animals. These surveillance systems are used to identify numerous diseases since they have several significant advantages: they cover a large part of the animal population and the costs associated with data collection and analysis are relatively low (Doherr and Audigé 2001; FAO 2014; Salman 2003). However, the performance of passive surveillance systems suffers from frequent under-reporting due to the lack of stakeholder awareness regarding a disease of interest. This may result in a failure to identify the disease (Hadorn et al. 2008) especially when it manifests in few or unspecific clinical signs (Doherr and Audigé 2001). In addition, potential fears of the disease’s consequences may also incite stakeholders to not report suspected cases (FAO 2014; Salman 2003). Under-reporting is especially problematic regarding the surveillance of exotic diseases, as exotic diseases have a low probability of occurrence, their symptoms usually not well known by practitioners, and the consequences of reporting an exotic disease in a new area may be dramatic.

3.3. Early warning in horse population

For the early detection of exotic VBDs in horses, both active and passive approaches could theoretically be implemented. However, as previously highlighted, one of the major drawbacks of the active approach is

that it can be very costly, especially when a disease is rare like exotic VBDs (Doherr and Audigé 2001). Risk-based active sampling could be implemented, but the equine industry suffers from a lack of accurate data regarding the populations at risk (e.g., details of animal movements, population size and location) which might complicate the planning and the implementation of such an active surveillance system. The efficiency of the classical passive surveillance approach in the early detection of an outbreak may also be limited due to the high probability of under-reporting, especially for exotic diseases.

Instead of relying on classical active or passive surveillance to detect new outbreaks, then, new approaches for estimating the probability of outbreak occurrence may constitute a promising way to improve the early detection of VBDs in horses. Different approaches can be considered in estimating this probability: classical risk assessment and syndromic surveillance.

4. Risk assessments

Risk assessments are the component of risk analysis that estimates the risks associated with a hazard, probability of hazard occurrence and its consequences (OIE 2010). Applied to exotic diseases, they are well-known tools for describing the probability of pathogen entry and spread within an area. The probability of entry is defined as the probability that a pathogen enters in a given area, considering all potential pathways of introduction and without considering the later steps of transmission (OIE 2014). The probability of “spread” is a vaguer concept and can include different sub-definitions, as presented by de Vos and colleagues (de Vos et al. 2011): (1) the probability of transmission, which is defined as the probability that the pathogen is able to spread to susceptible hosts in the area at risk, (2) the probability of establishment, which is the probability that the pathogen is able to spread to susceptible hosts and to susceptible vectors given the conditions of introduction, and (3) the probability of spread, which is the probability that the pathogen is able to spread in time and space, considering both local and long-distance dispersal. Regarding the specific issue of early detection, the probability of spread is irrelevant, as it is related more to the assessment of a disease’s impact when early surveillance has already failed to detect and control an outbreak. Conversely, the probabilities of transmission and establishment are especially interesting for early detection as they indicate the time period and the most suitable area for early spread of a pathogen. In particular, the probability of establishment, which takes into account the place and time of entry, is an interesting parameter with which to evaluate the likelihood of an infection actually leading to local spread.

The above approach gives a probability of outbreak occurrence based on risk factors such as the suitability of an environment and climate for disease transmission, or the presence of risky practices (e.g., importation of animals from infected area). It can be used by decision makers for risk mitigation and/or to enhance stake-holders’ awareness of rare or emerging diseases through risk maps, as has already been proposed for some endemic VBDs (e.g., surveillance of West Nile virus in California (Brown 2012)

(<http://www.westnile.ca.gov/>) or tick-borne diseases in Europe (Beugnet et al. 2009) (<http://www.fleatickrisk.com/FR/Pages/Home.aspx>)).

5. Syndromic surveillance

To enhance traditional passive surveillance systems, methods based on the analysis of pre-diagnostic and unspecific routinely collected data were developed at the beginning of the 21st century. Such approaches, referred to as syndromic surveillance, aim to identify the early, often weak, signal of an outbreak in the absence of an accurate identification of the disease by medical practitioners or laboratories. There is no single and commonly accepted definition for syndromic surveillance but it is commonly accepted that it focuses on data collected prior to clinical diagnosis or laboratory confirmation (Katz et al. 2011; Shmueli and Burkom 2010). First developed in human medicine, it is now also widely used in veterinary medicine (Dórea et al. 2011); indeed, a recent review by Dupuy and colleagues (Dupuy et al. 2013a) identified at least 27 syndromic surveillance systems or initiatives in 12 European countries. However, regarding horses, few syndromic surveillance initiatives are in place and only two have been explicitly identified: one in UK with Equine quarterly surveillance reports (DEFRA/AHS/BEVA 2015) and another another in The Netherlands with the GD monitor system (Rockx et al. 2006).

Syndromic surveillance provides a risk of outbreak occurrence based on the abnormal evolution of a health-related indicator. Such approaches can be used to rapidly detect a well-known disease or new pathogen without *a priori* consideration and they thus promise to strengthen surveillance of VBDs in horses. However, because they rely on health-related indicators, syndromic surveillance usually has a low specificity (variations in the indicator might be due to disease or to another event) and it is not able to take into account other epidemiological information available for a disease, such as environmental risk factors.

B. RESEARCH QUESTION

In the present work we explore various sources of information that shed light on the probability of occurrence of a newly introduced epidemic, focusing on exotic VBDs in horses. This is approached from various ends: is there a probability of entry? is there a probability of establishment? and is there a change in clinical signs, or other health-related indicator, that may signal such an epidemic?

The risk assessment for pathogen entry and establishment gives a risk profile of outbreak occurrence based on risk factors. However, an outbreak may also occur (with low probability) in a lower risk area instead of a higher risk area. The value of risk assessment for early detection is therefore more as supporting evidence than as formal evidence of an outbreak.

Similarly, syndromic surveillance gives a risk of outbreak occurrence based on the abnormal evolution of a health indicator. However, an outbreak may also occur without modification of such indicators. Even when significant modification is detected, the signal is often very unspecific and might occasionally be due to random fluctuation or to the effects of another outbreak or similar event. Again, then, the value of syndromic surveillance for early detection is more as supporting evidence than as formal evidence of an outbreak.

Risk assessments and syndromic surveillance can both suggest the possibility of a newly introduced epidemic, but they do not prove the presence of the disease. Using these approaches in concert, however, can increase the amount of evidence available and can be a way to improve confidence in predictions of newly introduced epidemics. In doing so, the interaction between risk assessments and syndromic surveillance must be considered carefully as these techniques are not fully independent. Indeed, for example, increased awareness via risk assessments might change the value of syndromic surveillance, because awareness will lead to increased reporting of data and thus to more numerous false alarms.

All these risk indicators are highly variable over time and space due to the unique nature of VBDs, which have a strong spatiotemporal pattern that is influenced by climatic and environmental factors (Altizer et al. 2006; Gage et al. 2008; Reisen 2010). Variability can also arise as a result of other non-biological fluctuations related to features of international trade, animal production, and so on. Spatial and temporal analyses must thus be performed in order to provide accurate data on the probability of occurrence of a newly introduced VBD.

To address the challenges of establishing early warning systems for VBDs in horses, the present work explores spatiotemporal risk assessments and syndromic surveillance, alone and in concert. In particular, we describe the limits and advantages of both methods in order to arrive at a new and more valuable approach for early warning systems.

C. CASES STUDIES

As case studies, we focus on three emerging vector-borne diseases found in French horses: African horse sickness, equine encephalosis, and West Nile fever.

1. French equine industry

1.1. Equine population

In France there are between 900,000 and 1,000,000 horses, which are mainly used for sport and leisure. The horse population, as estimated by IFCE-SIRE, is more concentrated in Basse-Normandie (10% of the population) (IFCE - les Haras nationaux 2011) (see Figure 1). However, the exact number of horses and their geographical locations are unknown.

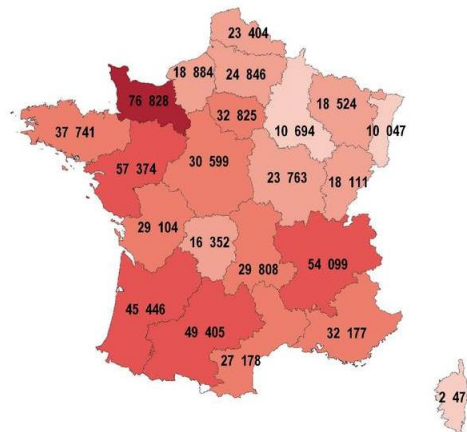


Figure 1: Estimated geographical distribution of horses population (IFCE-SIRE)

1.2. Organization of the equine industry

The equine industry in France is an important sector: it represents 4.3% of the full-time-job-equivalent in the agricultural sector and produces around 12 billion euros of revenue per year, mainly generated by horse-race betting operations (Lebrun 2010). Furthermore, France is the world's fourth-largest exporter of horses and its equine industry is growing, with an increase in both the number of breeders and the size of the breeding population since the last decade (IFCE - les Haras nationaux 2011).

Nonetheless, the French equine industry is fragmented, with different organizations in charge of the various sub-industries: racing, sport and leisure, breeding, and butchery. All of these sub-industries are further subdivided into several sectors, each with its own specific regulations and organizations as presented in **Erreur ! Source du renvoi introuvable.** The 'Institut français du cheval et de l'équitation'

(IFCE) is a central institute which manages the SIRE database (*'système d'identification des équidés'*) which collects all the data available on French microchipped equids.

Table 1: Main players in the French equine industry

Horse racing industry	<ul style="list-style-type: none"> - 'France galop': Grants authorization to train and ride horses for gallop racing; updates racing regulations and stud book organization - 'Cheval français': Grants authorization to train and ride horses for harness racing; updates racing regulations and stud book organization - 'Fédération nationale des courses françaises': Responsible for harmonization of racing regulations - 10 regional federations
Sport and leisure industry	<ul style="list-style-type: none"> - 'Fédération interprofessionnelle du cheval de sport, loisir travail': Cross-sector organization; promotes horse-related activities - 'Fédération nationale de cheval': Breeders' association - 'Association syndicale des étalonniers particuliers': Stallion raisers' association - 'Chambre syndicale du commerce des chevaux de France': Professional traders' association - 'Groupement hippique nationale': Riding centers' association - 'Syndicat National des Exploitants d'établissements Professionnels Enseignant l'équitation': Riding centers' association - 'Société hippique française': Coordinates genetic selection; riders' association - 'Fédération française d'équitation': Issues required licenses and complementary certificates; develops regulations for competitions - 'Acteurs et cavaliers de sauts d'obstacles en France': Riders' association - 'Haras nationaux': Confirms the origins of horses; promotes horse industry development
Butchery industry	<ul style="list-style-type: none"> - 'Interbev equin': Promotes horse meat, develops marketing and production rules - 'France trait': Union of nine French associations of draft horses; coordinates genetic selection
Breeding industry	<ul style="list-style-type: none"> - Different breed-specific stud books: coordinate horse selection and genetic improvement; promote purebred horses

1.3. Disease surveillance in French horses

The French surveillance system for equine diseases is mainly passive. The diseases for which mandatory reporting to the French ministry is in place are reported in Table 2. In addition to this classical passive reporting system, the French network for the surveillance of equine diseases, or 'RESPE' (<http://www.respe.net/>), collects declarations from veterinary practitioners registered as sentinels throughout France. RESPE, which was established in 1999, also issues alerts on equine diseases, such as information on diseases detected in French horses or in neighboring countries. More than 500 sentinel veterinarians are involved and cover 92 out of 96 French regions (see Figure 2). The veterinarians fill out a standardized questionnaire online and send standardized samples for laboratory diagnosis.

Table 2: Mandatory notifiable diseases to the French ministry. Category 1 = diseases of serious concern for public health and/or for the industry and that require preventive and control measures for the general interest, Category 2 = other diseases that also require preventive and control measures for the collective interest.

Category 1	Horse specific: equine infectious anemia, African horse sickness, western and eastern equine encephalitis viruses, Venezuelan equine encephalitis virus
	Non-horse specific: rabies, botulism, brucellosis, anthrax, Aujeszky's disease, tuberculosis, Japanese encephalitis, West Nile virus, vesicular stomatitis virus
Category 2	Equine viral arteritis, contagious equine metritis, glanders, trichinosis

Active testing of horses is performed by private partners prior to sale and by public partners prior to export for three diseases: equine viral arteritis, equine infectious anemia, and contagious equine metritis. Active surveillance is also performed by private and public partners on breeding stock for certain breeds and on all stallions used for semen collection. No further active surveillance system exists for the surveillance of equine diseases in France.

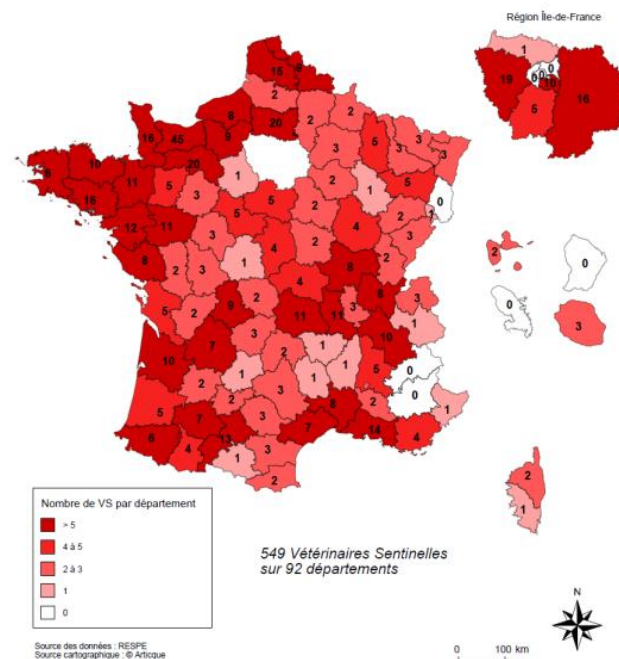


Figure 2: Geographic distribution of the French sentinel veterinarians involved in RESPE.

2. Diseases of interest

2.1. African horse sickness

Like the Bluetongue and Schmallenberg viruses, African horse sickness (AHS) is a *Culicoides*-borne disease, and it has recently been highlighted as a potential threat for Europe (Zimmerli et al. 2010)(MacLachlan and Guthrie 2010). The disease is caused by a virus belonging to the *Orbivirus* genus of the Reoviridae family (Mellor and Hamblin 2004). There are nine different serotypes that confer some degree of cross-protective immunity (Mellor and Hamblin 2004). The virus is considered endemic in sub-Saharan countries, with rare outbreaks in North Africa and western Asia (MacLachlan and Guthrie 2010). The last AHS outbreak in Europe occurred between 1987 and 1990 in the Iberian Peninsula and resulted from the importation of infected zebras (Rodriguez et al. 1992a).

AHS is a non-zoonotic disease that affects all extant Equidae, although morbidity and mortality vary among species: as many as 90% of infected horses die within one week, while infection is largely subclinical in zebras (Mellor and Hamblin 2004; Wilson et al. 2009). It is considered to be one of the most devastating diseases that affect equids. Four forms of the disease exist: horse sickness fever (moderate

fever and no mortality), cardiac form (subcutaneous edema, particularly of the head, neck, chest, and supraorbital fossae, mortality 50%), mixed form (combination of the cardiac and pulmonary forms, mortality 70%) and pulmonary form (sudden death, severe dyspnea, mortality 95%) (Mellor and Hamblin 2004). The incubation time is from 3 to 15 days (Theiler 1910).

2.2. *Equine encephalosis*

Equine encephalosis (EE) is caused by a virus of the *Orbivirus* genus of the Reoviridae family, and encompasses seven different serotypes (Dhama et al. 2014; Viljoen and Huismans 1989). Similarly to AHS, EE has been recently highlighted as a potential threat for Europe (MacLachlan and Guthrie 2010; Zimmerli et al. 2010). Indeed, AHS and EE viruses are similar in many aspects: both are non-zoonotic *Culicoides*-borne members of genus *Orbivirus* that share the same vectors (Venter et al. 2002; Venter et al. 2000) and more or less the same geographical distribution. Like AHS, EE is also considered to be endemic in sub-Saharan countries, with rare outbreaks in North Africa and western Asia (Mildenberg et al. 2009; Wescott et al. 2013). The last major outbreak was reported in Israel in 2009 (Mildenberg et al. 2009). EE has never been observed in Europe (Dhama et al. 2014).

The epidemiology of EE is similar to AHS but the pathogenicities of the two viruses are different. The incubation time of EE is shorter (2-6 days (Theiler 1910)) and its transmission rate is higher (Lord et al. 2002). Moreover, despite the fact that EE was initially described as a “fever in horses simulating horse-sickness” (Theiler 1910), the symptoms of the two diseases are different. In particular, the mortality rate of EE is always low. In contrast to AHS, which causes severe cardiac and pulmonary symptoms, EE is characterized by a wide range of symptoms, such as abortions during the first 5-6 months of gestation, respiratory signs (e.g., nasal discharge, cough), and encephalitis (Dhama et al. 2014).

2.3. *West Nile virus*

West Nile virus (WNV) is a mosquito-borne arbovirus belonging to the genus *Flavivirus* (family Flaviviridae) and mainly transmitted by mosquitoes from the genus *Culex* (family Culicidae). Two distinct lineages exist: lineage 1 causes outbreaks throughout the world, while lineage 2 was limited to Africa until 2008, when it was introduced to Europe. Since the discovery of WNV in 1937 in Uganda (Smithburn et al. 1940), the geographic distribution of the virus has expanded and the disease is now considered endemic in Africa, Asia, Europe, Australia, the Caribbean, and the Americas (Campbell et al. 2002; Ozdenerol et al. 2013). In Europe, WNV emerged in the 1960s and several outbreaks have subsequently been documented in many European countries (Calistri et al. 2010). Even if the virus is now endemic in large parts of Europe, the number of reported outbreaks is presently increasing in Southern and Eastern Europe (e.g., Italy, Greece, Bulgaria, Croatia, Serbia, Albania) (Di Sabatino et al. 2014). This increase in the number of outbreaks, combined with the recent introduction and spread of lineage 2 in Europe (which has been associated with

severe cases in humans, horses, and birds (Bakonyi et al. 2006; Calzolari et al. 2013; Hernández-Triana et al. 2014)), contribute to the growing concern about WNV in Europe.

The enzootic cycle of WNV is driven by its continuous transmission to susceptible bird species through adult mosquitoes. Its main hosts are birds, but the virus also affects more than 30 non-avian species. The susceptibilities of birds to WNV infection differ, with those in the order Passeriformes being most susceptible, followed by birds in the order Charadriiformes and domestic geese (order Anseriformes). Psittacine and gallinaceous birds are less susceptible. WNV in birds is usually asymptomatic, but may cause nonspecific clinical signs, neurological signs, and death (Pérez-Ramírez et al. 2014; Steele et al. 2000). Of non-avian species, the most affected are humans and horses (Kramer and Bernard 2001; Van der Meulen et al. 2005). In horses the clinical signs of WNV are almost exclusively neurological and reflect its pathology in the central nervous system (Cantile et al. 2000; Castillo-Olivares and Wood 2004). In humans two presentations of the disease are reported: uncomplicated WN fever (headache and myalgia, often accompanied by gastrointestinal symptoms) and WN meningoencephalitis (typical meningitis or encephalitis) (Campbell et al. 2002; Colpitts et al. 2012).

D. OUTLINES OF THE WORK

Chapter 1 presents the context of this work and the associated research question. Chapter 2 gives first a brief presentation of the methods used for spatiotemporal quantitative risk assessments of the entry and establishment of VBDs in France, then presents spatiotemporal risk analyses that were performed to assess the probabilities of AHS and EE entry and establishment. Two routes of viral entry were considered together and two methods were used to assess the probability of viral establishment.

Chapter 3 first provides an overview on current approaches in syndromic surveillance. Then, an application of a classical method is presented which explores the impact of pre-processing methods on surveillance system performance. Finally, we discuss the application of Bayes' rules to syndromic surveillance with the goal of generating a quantitative output from syndromic surveillance and combining this with other epidemiological information.

Chapter 4 presents the combination of risk assessments with syndromic surveillance data using a Bayesian approach. We present a combination of various sources of epidemiological information, which originate from different syndromic surveillance systems and/or from syndromic surveillance and risk analysis.

Chapter 5 concludes with a discussion of the reliability and transparency of these complex surveillance systems and their usefulness in supporting decision-making.

CHAPTER II: QUANTITATIVE RISK ASSESSMENTS

This chapter explores quantitative risk assessments as a way to assess the risk of potential newly introduced VBDs in horses. Early detection of an exotic vector-borne pathogen can reduce the impact of the disease. It relies on two factors: the probability of pathogen entry and the probability of pathogen transmission and establishment. In fact, the probability of spatial and numeric spread is irrelevant to this process, as it is related more to the assessment of a disease's impact after early detection has already failed to detect and prevent the outbreak. The parameters needed to assess these probabilities are specific to each VBD, given that a vector's biology is closely linked to season and environment. A brief overview of the methods used for quantitative risk assessments of a VBD's entry, transmission, and establishment is presented in Chapter II.A. This review highlighted that combination of probability of entry, probability of transmission and probability of establishment is still rarely implemented in animal health, as is a thorough study of multiple routes of pathogen entry. Building on this review, we thus developed a quantitative model to assess the probability of entry and establishment of AHS, one of the most devastating equine diseases known (Chapter II.B). We implemented spatiotemporal analysis to take into account the close link between VBDs and season/environment, and to consider other non-biological fluctuations related to features of international trade and animal production. Then, we evaluated the feasibility of adapting this model to other VBDs in horses (Chapter II.C). For that purpose, we focused on equine encephalosis (EE), as this disease is similar to AHS. The respective probabilities of entry into France were compared for each disease. Finally, we discuss the advantages and drawbacks of our model in assessing the risks of newly introduced VBDs in horses (Chapter II.D).

A. OVERVIEW

This section aims to present the general principle of risk assessment and explain how the probabilities of pathogen entry, transmission, and establishment can be assessed in the specific context of VBDs. We restrict ourselves to quantitative risk assessment and to the most popular methods for risk estimation.

1. General principle of risk assessment

1.1. Definitions and objectives

Risk assessment is the component of risk analysis that estimates the risk associated with a hazard (OIE 2014). The OIE has defined a hazard as a “biological, chemical, or physical agent in, or condition of, an animal or animal product with the potential to cause an adverse health effect” and a risk as “the likelihood of the occurrence and the likely magnitude of the biological and economic consequences of an adverse event or effect to animal or human health”.

The purpose of risk assessment is not so much to predict the introduction of a hazard but rather to help managers to better understand the associated risks (e.g., relative contributions of various factors, current areas of distribution, pathways for introduction, effectiveness of risk prevention actions). Risk assessment is clearly separated from risk management, but actively collaborates to achieve the ultimate goal: implement measures that ensure the appropriate level of protection (Giovannini et al. 2004; OIE 2014).

Risk assessment may focus on different components, such as assessment of the entry, exposure, or consequences of a VBD, or on an estimation of the total risk by combining the results of these separate components in an overall assessment. To our knowledge, only a few papers in the animal health literature have linked these different probabilities in a quantitative manner: in a literature search, only three papers were found that addressed a combination of the probability of entry of a virus and its probability of establishment (EFSA 2009; Napp et al. 2012; de Vos et al. 2012). Similarly, a comprehensive approach to the routes of entry is rarely used; we found only one paper that analyzed a combination of routes of virus entry (Kilpatrick et al. 2004).

The entry assessment is conducted using the probability that the pathogen of interest enters the area at risk via any possible pathway, without regard to later steps of transmission.

The exposure assessment includes analysis of various factors, as presented by de Vos and colleagues (de Vos et al. 2011): (1) the probability of transmission, which is defined as the probability that the pathogen is able to spread to susceptible hosts in the area at risk, (2) the probability of establishment, which is the probability that the pathogen is able to spread to susceptible hosts and to susceptible vectors (and vice versa) given the conditions of introduction, and (3) the probability of extended spread, which is the

probability that the pathogen is able to spread in time and space, considering both local and long-distance dispersal.

The consequence assessment rates the impact of the disease and includes health, economic, social, ethical, and environmental considerations (Vose 2008). Disease persistence to next season won't be considered in this work.

1.2. Method

The preliminary step of any risk assessment is to identify the hazard of interest. Hazard identification is fundamental in defining the objective of the risk assessment and must be carefully implemented (Vose 2008).

To perform risk assessment, several guidelines are available (see for example (Codex Alimentarius 1999; EFSA 2010; OIE 2014; USDA 2012; de Vos et al. 2011)) and no single method is applicable to all risk assessments. However, the following principles identified by OIE in the International Animal Health Code (OIE 2014) should ensure the quality of risk assessment:

- Risk assessment should be flexible enough to deal with the complexity of real-life situations.
- The risk assessment should be based on the best available information that is in accord with current scientific thinking. The assessment should be well-documented and supported with references to the scientific literature and other sources, including expert opinion.
- Consistency in risk assessment methods should be encouraged and transparency is essential in order to ensure fairness and objectivity, consistency in decision making, and ease of understanding by all interested parties.
- Risk assessments should document the uncertainties, the assumptions made, and the effect of these on the final risk estimate.
- The risk assessment should be amenable to updates when additional information becomes available.

Risk assessments can be either qualitative or quantitative, and both approaches are valuable. Qualitative assessments describe the risk in words (e.g., low, moderate, high), whereas quantitative assessments express the risk in numeric terms. Qualitative assessments are performed when no proper evaluation of the order of magnitude of uncertainties is possible while these uncertainties are typically high. Since there is no quantitative evaluation, there is also no proper mathematical model. Qualitative assessments are often used for routine decision making (OIE 2014). However, qualitative assessments do not provide sufficient information to accurately discriminate between small and large risks (Cox et al. 2005).

Quantitative assessments provide more detailed information on the risk and can be more useful in distinguishing periods and areas at higher risk. However, the performance of a quantitative risk analysis is limited by the quality of data available. In addition, quantitative risk assessment can be deterministic or

stochastic. Deterministic approaches produce a single outcome from a given set of parameters, uncertainty can be included but stochastic effects are usually ignored or crudely estimated. Stochastic approaches directly calculate the risk while also taking into account uncertainty and/or variability due to stochastic variation in input parameters. They produce a probability distribution of possible outcomes distinguishing impact via uncertainty and stochasticity.

2. Probability of entry

2.1. Definition

The probability of entry was previously referred to as the probability of release (OIE 2010). In 2014, the new version of the OIE's Terrestrial Animal Health Code (OIE 2014) specified this a bit further and adopted the following definition: "Entry assessment consists of describing the biological pathway(s) necessary for an importation activity to introduce pathogenic agents into a particular environment, and estimating the probability of that complete process occurring".

2.2. Estimation for VBDs

2.2.1. Routes of entry

The first step of the assessment of pathogen entry is to identify the potential routes available for introduction of the pathogen. Considering the specific case of VBDs, de Vos and colleagues (de Vos et al. 2011) listed the following pathways which should be considered in assessing the probability of VBD entry:

- Entry of infected live animals via importation of livestock, zoo animals, pets, wildlife, or migratory birds;
- Entry of an infected vector or its eggs or larvae through expansion of the vector's habitat or transport with wind, tires, plant materials, transport vehicles, animals, humans, manure, or soil;
- Import of contaminated biological material, such as semen, ova, embryos, serum, plasma, and modified live vaccines;
- Import of contaminated animal products such as meat, milk, eggs, bush meat, and animal by-products (feathers, animal proteins, animal fats);
- Entry of infected humans.

The importation of animals and transport of materials or other products can be legal or illegal. Illegal imports could contribute substantially to the probability of entry, but are obviously difficult to estimate and to quantify, and will therefore not be discussed further in this work (Chaber et al. 2010; Hartnett et al. 2007; Smith et al. 2012).

2.2.2. Calculation

The scenario tree, or scenario pathway, approach is commonly used to estimate the probability of pathogen entry (Vose 2008). Several examples are available for vector-borne pathogens, in particular for Rift Valley fever (Abdo-Salem et al. 2011), West Nile fever (Bessell et al. 2014; Douglas et al. 2007; Kilpatrick et al. 2004), equine infectious anemia (Asseged et al. 2012), and bluetongue (Hoar et al. 2004). Other examples are also available for non-vector-borne pathogens like classical swine fever (Bronsvort et al. 2008) and foot-and-mouth disease (Yu et al. 1997).

The principle is first to describe all consecutive steps that result in disease entry. The complexity of these steps can vary according to the routes of entry considered (de Vos et al. 2011). For example, when an infected animal is legally imported, the animal must be viraemic (or latently infected) and the infection not detected during import procedure. However, when an infected wild bird enters the country via migratory routes, the only step to be considered in the probability of pathogen entry is that the animal is viraemic or latently infected. Each step has a conditional probability of occurrence and the probability of entry is calculated by multiplying the probabilities of all steps along the tree.

Model calculations can be used to quantitatively combine all these probabilities. These are especially useful when there is uncertainty and/or variability due to the presence of stochastic parameters, which is often the case in risk analysis.

3. Probabilities of transmission and establishment

3.1. Probability of transmission

3.1.1. Definition

The probability of transmission can be evaluated by calculating the basic reproductive number (R_0) (see for example (de Koeijer et al. 2002) and (Fischer et al. 2013)), which represents the expected number of secondary cases produced, in a completely susceptible population, by a “typical” infected individual during its entire period of infectiousness (Diekmann et al. 1990). The probability of transmission is thus not a probability but the number of new cases generated from one initial infected case.

If $R_0 > 1$, then, on average, each infected animal will generate more than one infected animal, and the infection will increase exponentially. If $R_0 < 1$, an infected animal is not able to infect more than one other animal, on average, and the infection should die out.

Due to the ease of interpretation of R_0 , it is very useful in distinguishing at-risk situations from those not at risk. This is one of the reasons why it has become an important parameter in risk assessment. In particular,

R_0 is often used to assess the transmission of exotic infectious diseases for which, by definition, the only infected individual in a population is the introduced one.

3.1.2. Estimation for VBDs

Numerous approaches are available to estimate R_0 , but with vector-borne infection, the estimation of R_0 is slightly more complicated than for diseases with direct transmission. Indeed, with VBDs there is not only one population of infected/susceptible individuals to be considered, but at least two: hosts and vectors. Thus two transmission steps must be assessed: one infectious host with a fully susceptible vector population, and one infectious vector with a fully susceptible host population.

The following parameters must be considered to assess the probability of VBD transmission (de Vos et al. 2011):

- Host density;
- Vector abundance;
- Biting rate;
- Transmission probability per bite from host to vector and from vector to host;
- Vector biology (e.g., survival rate, extrinsic incubation period).

Moreover, the vector-related data are highly dependent on environmental and climatic factors. Thus the R_0 for VBDs also depends on the time period and the geographical location in which the pathogen enters the area at risk.

Classical models for the assessment of VBD transmission probabilities are based on systems with one host and one vector or with two hosts and one vector. More details can be found, for example, in studies conducted on malaria (Macdonald 1955; Ponçon et al. 2008), African horse sickness (Backer and Nodelijk 2011; Lord et al. 1996), Rift valley fever (Fischer et al. 2013), and bluetongue virus (Brugger and Rubel 2013; Gubbins et al. 2008; Hartemink et al. 2009; de Koeijer et al. 2011). However, other authors have also proposed approaches based on two-host, two-vector models (Turner et al. 2013).

3.2. Probability of establishment

3.2.1. Definition

The probability of establishment represents what happens in terms of pathogen survival and growth during a transition phase between a pathogen's entry and its actual spread. Establishment can be considered the initial spread of the disease and is expressed as the probability that "the infection has passed from a host via a vector to an indigenous host, while the basic reproduction number, R_0 , is higher than 1" (de Vos et al. 2011).

First, the probability of establishment depends on the route of pathogen entry. Beyond this, it depends on environmental and climatic conditions and therefore on the time period and the location in which the pathogen enters the area at risk.

3.2.2. Estimation for VBDs

Methods such as those already presented in the previous section on the probability of entry assessment (i.e. the scenario tree method) can be used to estimate the probability of establishment. As previously presented, the principle is to describe all consecutive steps that result in disease establishment.

For VBDs, the following steps must be taken into particular consideration (*de Vos et al. 2011*):

- Pathway for introduction: route of exposure of indigenous host or vector;
- Time of disease entry: temperature, humidity, and other parameters that may influence pathogen transmission;
- Geographic location of disease entry: host density, vector abundance;
- Vector-host interaction: biting rate, transmission probability per bite from vector to host and from host to vector.

Studies conducted on African horse sickness (*de Vos et al. 2012*), bluetongue (*Napp et al. 2012*), and epizootic hemorrhagic disease (*EFSA 2009*) provide examples of the assessment of the probability of VBD establishment using different routes of pathogen introduction (e.g., infectious host or infectious vector).

4. Conclusion

Quantitative risk assessments are common methods used to assess the risk posed by exotic pathogens. Various methods are available to implement risk assessment. Notably, different components of risk assessment (i.e. probabilities of entry, transmission, and establishment) can be assessed independently or together according to the objective of the risk assessment. Considering the case of VBDs, specific parameters have to be taken into account to deal with the two steps of virus transmission (from vector to host and from host to vector). In particular, the time and the area of pathogen entry are important in assessments of the probabilities of pathogen transmission and establishment as a vector's biology is closely linked to season and environment. However, analyses that combine the probabilities of entry, transmission, and establishment are still rarely implemented in animal health; the same is true regarding studies of combinations of different routes of entry. The development of such models could be a way to better understand and manage the risk associated with a disease.

B. PROBABILITIES OF ENTRY AND ESTABLISHMENT

1. Introduction

In this section we present a spatiotemporal assessment of the probability of introduction of African horse sickness (AHS) to France. We developed a model that combined the probability of entry with the probability of establishment and included two routes of virus entry, expecting that such a model would enable a greater understanding of the risk associated with AHS in France compared to risk assessments involving only a single probability parameter or a single route of introduction.

This work was published in BMC Vet Research (see below for the main text and Appendix 1, Appendix 2, and Appendix 3 for the additional files of the paper), two professional publications in French journals (see Appendix and Appendix 5), and one poster presentations at the EPIZONE meeting in 2013 (see Appendix 6).

RESEARCH ARTICLE

Open Access

A spatiotemporal model to assess the introduction risk of African horse sickness by import of animals and vectors in France

C. Faverjon^{1*}, A. Leblond², P. Hendrikx³, T. Balenghien⁴, C. J. de Vos⁵, E.A.J. Fischer⁵ and A.A. de Koeijer⁵

Abstract

Background: African horse sickness (AHS) is a major, *Culicoides*-borne viral disease in equines whose introduction into Europe could have dramatic consequences. The disease is considered to be endemic in sub-Saharan Africa. Recent introductions of other *Culicoides*-borne viruses (bluetongue and Schmallenberg) into northern Europe have highlighted the risk that AHS may arrive in Europe as well. The aim of our study was to provide a spatiotemporal quantitative risk model of AHS introduction into France. The study focused on two pathways of introduction: the arrival of an infectious host (PW-host) and the arrival of an infectious *Culicoides* midge via the livestock trade (PW-vector). The risk of introduction was calculated by determining the probability of an infectious animal or vector entering the country and the probability of the virus then becoming established: i.e., the virus's arrival in France resulting in at least one local equine host being infected by one local vector. This risk was assessed using data from three consecutive years (2010 to 2012) for 22 regions in France.

Results: The results of the model indicate that the annual risk of AHS being introduced to France is very low but that major spatiotemporal differences exist. For both introduction pathways, risk is higher from July to October and peaks in July. In general, regions with warmer climates are more at risk, as are colder regions with larger equine populations; however, regional variation in animal importation patterns (number and species) also play a major role in determining risk. Despite the low probability that AHSV is present in the EU, intra-EU trade of equines contributes most to the risk of AHSV introduction to France because it involves a large number of horse movements.

Conclusion: It is important to address spatiotemporal differences when assessing the risk of AHS introduction and thus also when implementing efficient surveillance efforts. The methods and results of this study may help develop surveillance techniques and other risk reduction measures that will prevent the introduction of AHS or minimize AHS' potential impact once introduced, both in France and the rest of Europe.

Keywords: African horse sickness, Equine movements, Import risk assessment, Risk of introduction, *Culicoides*, Quantitative risk, Midge

Background

African Horse Sickness (AHS) is a highly fatal viral vector-borne disease that is transmitted among equine hosts by *Culicoides* midges (Diptera: Ceratoponidae) [1, 2]. It affects all extant Equidae, but morbidity and mortality vary among species: as many as 90 % of infected horses die within one week, while infection is largely subclinical in zebras [3, 4]. AHS virus (AHSV) is an orbivirus, and there

are nine different AHSV serotypes that confer some degree of cross-protective immunity [4]. AHS is considered to be endemic in sub-Saharan Africa, where the zebra acts as a reservoir [5]. Rare outbreaks have occurred in North Africa, western Asia, and the Iberian Peninsula, where they have persisted for only a few years [6]. The last outbreak in Europe occurred in the Iberian Peninsula, between 1987 and 1990, and caused the death of more than 1,350 horses, either directly or as a result of control measures [7].

The recent introduction into northern Europe of bluetongue virus (specifically BTV-8, in 2006 [8]) and Schmallenberg virus (in 2011 [9]), both transmitted by

* Correspondence: celine.faverjon@vetagro-sup.fr

¹INRA UR346 Animal Epidemiology, Vetagrosup, F-69280 Marcy l'Etoile, France

Full list of author information is available at the end of the article



Culicoides midges, highlights the relevance of assessing the risk that AHSV will be introduced to Europe [10]. It is particularly crucial to conduct a risk assessment analysis for France, as the country encompasses different ecosystems, including Mediterranean zones, where *Culicoides imicola*—considered to be the major vector of AHSV worldwide—is very abundant, and non-Mediterranean temperate zones, where *Culicoides obsoletus*—a potential AHSV vector—is dominant [11, 1, 12]. Moreover, France contains between 900,000 and 1 million equines, is the world's 4th largest exporter of horses, and has a horse industry that produces around 12 billion euros of revenue per year [13]. If AHS arrived in France, it could have devastating consequences, similar to those predicted for other EU members such as the United Kingdom (UK) [14], Ireland [15], and the Netherlands [16].

Introduction risks have recently been quantitatively assessed for similar vector-borne diseases, such as BTV [17–19], West Nile Virus [20, 21] and eastern and western equine encephalomyelitis [22], Venezuelan equine encephalitis [22], and Japanese encephalitis [22]. However, these studies mostly took into account only one pathway of introduction; different introduction pathways have rarely been examined in tandem. To explore AHSV in particular, a qualitative risk assessment analysis that accounted for multiple pathways of introduction was conducted in the UK [23]; the results suggested that the most likely pathway of introduction would be the arrival of an infectious host. This pathway of introduction was also examined by a quantitative risk assessment analysis of the likelihood that AHSV would be introduced to the Netherlands [24]. As AHSV is closely related to BTV and shares the same vectors, information on BTV introduction pathways could be helpful when assessing the risk that AHSV will be introduced to Europe. Several studies have indicated that long-distance, wind-mediated transport of *Culicoides* might have played a role in the introduction and spread of several BTV strains in Europe [25–27]. In particular, studies aimed at understanding the introduction of BTV-8 have indicated that the legal importation of an infectious host is unlikely to have caused the epidemic observed in 2006 [28, 29]. It is thought that the risk presented by other pathways of introduction, such as the introduction of a single *Culicoides* midge through intracontinental transport and trade networks [19], is low. Integrative studies are required to quantify and combine information on different pathways of introduction to better understand and confront the risks posed by vector-borne diseases [30].

In this study, we performed a quantitative risk assessment analysis of the introduction of AHSV to France. We focused on two pathways of introduction: the arrival of an infectious host and the arrival of an infectious

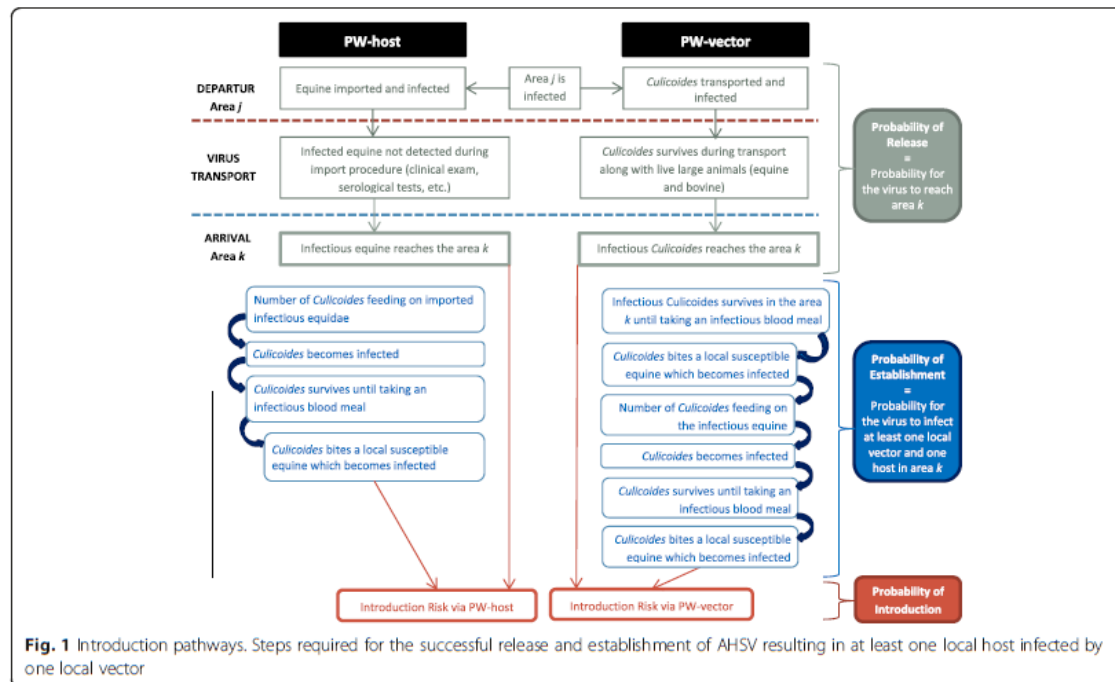
Culicoides midge via the livestock trade. Introduction is defined here as the probability that an infectious host or vector will be released in such a way that at least one local host ends up infected by one local vector (establishment). The subsequent spread of the disease is not examined. As the initial infection of a local host will depend on spatial (e.g., the number of local hosts) and temporal factors (e.g., seasonal vector abundance), the probability of establishment will vary depending on location and time period. The objective of this study was to quantify the risk of introduction associated with a given time period and region for the two pathways of introduction under consideration, which could thus offer insight into temporal and regional variation in introduction risk. Furthermore, evaluating these two pathways of introduction could help optimize risk-mitigating control and surveillance measures.

Methods

Risk associated with introduction pathways and initial assumptions

To quantify the risk of AHSV introduction associated with the two introduction pathways, risk assessment analysis was conducted using the framework developed by de Vos et al. [31]. Although other potential pathways of introduction exist [23], we restricted ourselves to the two most probable: the arrival of an infectious equine and the arrival of an infectious *Culicoides* midge. Only the legal, registered horse trade was taken into account because no data exist on the illegal horse trade. In the analysis, only the introduction of an adult vector was considered since transovarial transmission of the virus has not been observed in *Culicoides* [1, 11, 15]. *Culicoides* midges are hematophagous and tend to stay close to their mammalian hosts (mainly large mammals [32, 33]). They are rarely found in vehicles of transport (such as aircraft or trucks) or merchandise when insect surveys are conducted [21, 34]. It is also uncommon to find *Culicoides* associated with plants or plant material [29, 35]. As a result, only *Culicoides* entering the country via the livestock trade were included in the analysis. Hence, the two main pathways of AHSV introduction examined in this study were: the legal importation of infectious equines (PW-host) and the arrival of infectious *Culicoides* as a consequence of livestock trade (PW-vector).

An introduction pathway was constructed to detail all the steps required for AHSV to be successfully released and become established in France (Fig. 1). This introduction pathway was evaluated using a stochastic risk simulation model. Monthly introduction probabilities were calculated using data from three consecutive years (2010 to 2012) for each area of arrival within France. A total of 22 such areas were defined.



Sources of risk

First, the world's countries were grouped into three categories as per De Vos et al. [24]: (1) high-risk regions where the disease is considered to be endemic; (2) low-risk regions that have experienced AHS outbreaks in the past and/or where the main vector, *C. imicola*, is present; and (3) very-low-risk regions (all other countries). Since the main vector is not present in very-low-risk regions, we assumed that it is very unlikely that they would produce exports containing infectious vectors; consequently, very-low-risk regions were ignored for this pathway of introduction. In addition, because EU regulations differ for imports arriving from EU versus non-EU countries [36], we also distinguished between (a) EU members and (b) non-EU members. Five departure regions were thus defined (Fig. 2). Imports from non-EU countries were placed in one of two categories based on their point of arrival in Europe: whether they were shipped directly to France or whether they arrived via another EU country, because animals stopping in another country were considered to experience longer traveling times. Furthermore, equine imports were grouped according to species: (1) horses; (2) donkeys, mules, and hinnies; and (3) zebras.

The model

PW-host: introduction via an infectious host

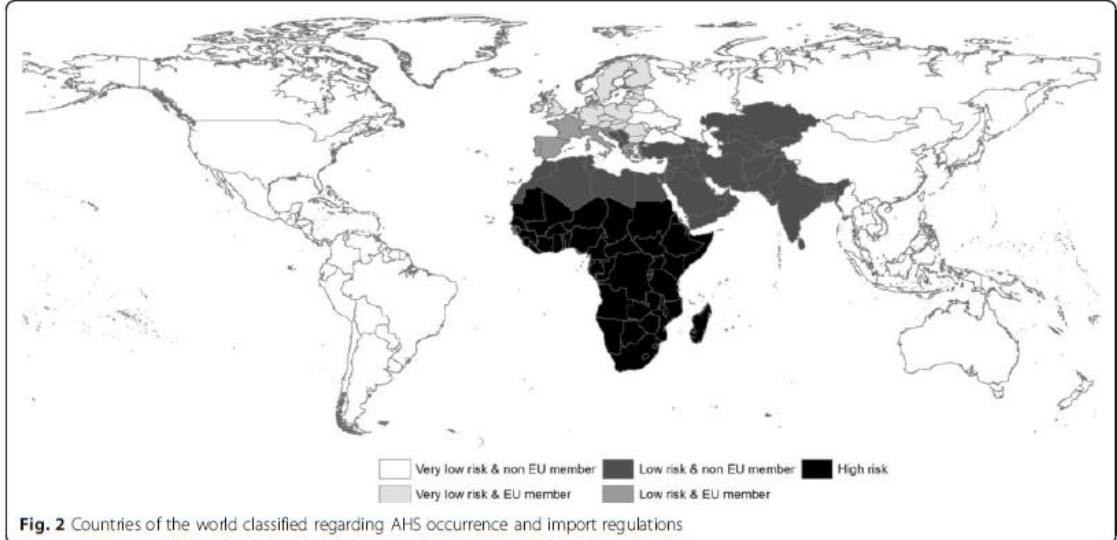
The probability of AHSV being introduced by species *i* from region *j* to area *k* in month *m* via an infectious

host (PW-host), ($introH_{ijkm}$), was defined as the probability of at least one infectious host of species *i* from region *j* arriving in area *k* in month *m* and of this arrival being followed by virus transmission to a local vector and host. This overall probability was defined as:

$$P(introH_{ijkm}) = 1 - [1 - P(relH_{ijkm}) \times P(estH_{ijkm})]^{eq_{ijkm}} \quad (1)$$

Where $P(relH_{ijkm})$ is the probability of an infectious equine of species *i* from region *j* being released in area *k* in month *m*; $P(estH_{ijkm})$ is the probability of an infection becoming established in month *m* given the release of one infectious equine of species *i* from region *j* in region *k*; and eq_{ijkm} is the number of equines of species *i* imported from region *j* arriving in area *k* in month *m*.

Release probabilities, $P(relH_{ijkm})$, were species specific since virus prevalence is different in different equines across the areas of origin, and different species show differences in their susceptibility to the disease. For instance, the release probability of horses is lower because horses have a shorter viremic period than do donkeys and zebras. $P(relH_{ijkm})$ also depended on the moment *z* of infection; the protective measures implemented before embarkation [36, 37]; and the duration of transport from region *j* to area *k* (t_{jk}). For imports coming from non-EU countries, t_{jk} was assumed to equal 1 day for



animals coming directly to France (mainly air travel) and 2 days for animals arriving via another EU country (initial air travel followed by land transport or subsequent air travel). For intra-EU trade, t_{jk} was assumed to be between 1–2 days (uniform distribution) because air and land transport are supposed to be used with equal frequency and France is assumed to be a maximum of 2 days away from everywhere else in EU [38]. $P(\text{rel}H_{ijkm})$ was defined as :

$$P(\text{rel}H_{ijkm}) = \frac{\sum_{z=1}^w [(\text{length period } z) \times P(\text{rel}H_{ijkmz})]}{\sum_{z=1}^w (\text{length period } z)} \quad (2)$$

where $P(\text{rel}H_{ijkmz})$ is the probability of release for an equine i infected during z . For a given region j , there were a total of w different time periods z during which an equine could become infected, depending on the importation procedures implemented for region j (e.g., before quarantine or during quarantine but before the first serological test, etc.). If j was a high-risk region, then $P(\text{rel}H_{ijkmz})$ was defined as:

$$P(\text{rel}H_{ijkmz}) = P(\text{inf}_{ijmz}) \times P(\text{vir}_{ijmz}) \times (1 - P(\text{CF1}_{iz})) \times (1 - P(\text{CF2}_{iz})) \times (1 - P(\text{clin}_{ijmz})) \times (1 - P(\text{trans}_{ijkz})) \quad (2.\text{bis})$$

Where $P(\text{inf}_{ijmz})$ is the probability of equine i being infected during time period z ; $P(\text{vir}_{ijmz})$ is the probability

of equine i infected during z becoming viremic after transport; $(1 - P(\text{CF1}_{iz}))$ and $(1 - P(\text{CF2}_{iz}))$ are the probabilities of equine i infected during z not being detected by the first and second serological tests, respectively; $(1 - P(\text{clin}_{ijmz}))$ is the probability of equine i infected during z not being detected by the clinical exam; and $(1 - P(\text{trans}_{ijkz}))$ is the probability of equine i infected during z not being detected during transport.

The probability of establishment via PW-host was defined as:

$$P(\text{est}H_{ijkm}) = 1 - [1 - \lambda_{HV} \times P(\text{surv}_{km}) \times b_{equik} \times \lambda_{VH}]^{cid_{ikm}} \quad (3)$$

where λ_{HV} and λ_{VH} are, respectively, the probabilities of a vector feeding on an infectious host becoming infected and of a host bitten by an infectious vector becoming infected; $P(\text{surv}_{km})$ is the probability of an infected *Culicoides* midge surviving until its first infectious blood meal; c_{ulikm} is the number of vectors feeding on an infectious imported host; and b_{equik} is the probability of a *Culicoides* midge biting an equine in area k . b_{equik} depends on the vector's preference for equines as hosts and on the cattle-to-equine ratio in area k .

The overall national and annual median probabilities were calculated based on the monthly regional values. The monthly national median probability of introduction was thus defined as:

$$P(\text{intro}H_{ijm}) = 1 - \prod_{K=1}^{22} (1 - P(\text{intro}H_{ijkm})) \quad (4)$$

and the annual national median probability of introduction was defined as:

$$P(\text{intro}H_{ij}) = 1 - \prod_{m=1}^{12} (1 - P(\text{intro}H_{ijm})) \quad (5)$$

The same formulas were used to define the probabilities of release and establishment at the national and annual levels. Using these formulas, the extreme values of these probabilities were determined by using the estimated the 5th and 95th percentiles for each region and month.

For a more detailed description of the probabilities and parameters used, see Additional files 1 and 2.

PW-vector: introduction via an infectious vector

Few data were found on the number of *Culicoides* midges transported with livestock (equines and cattle) over the distances of interest here; as a consequence, we assumed that the numbers could not be very high without having spurred notice and thus calculated the risk of release assuming that one *Culicoides* was transported with each animal. We defined the probability of establishment as the probability that this single vector was able to cause the infection of at least one local equine host by one local vector. $P(\text{intro}V_{jkm})$ was thus define as:

$$P(\text{intro}V_{jkm}) = 1 - [1 - P(\text{rel}V_{jkm}) \times P(\text{est}V_{jkm})]^{n_{jkm}} \quad (6)$$

where $P(\text{rel}V_{jkm})$ is the probability of a single infected *Culicoides* from region j being released in region k in month m ; $P(\text{est}V_{jkm})$ is the probability of establishment; and n_{jkm} is the number of livestock (equines and bovines) transported from region j to area k during month m .

The probability of release was defined as:

$$P(\text{rel}V_{jkm}) = P(\text{inf_culi}_{jm}) \times P(\text{trans}_{culi}_{jm}) \times P(\text{surv}_{trans}_{jkm}) \quad (7)$$

where $P(\text{inf_culi}_{jm})$ is the probability of a vector in region j becoming infected in month m ; $P(\text{trans}_{culi}_{jm})$ is the probability of a vector being transported post infection; and $P(\text{surv}_{trans}_{jkm})$ is the probability of a *Culicoides* surviving transport from region j to region k . $P(\text{surv}_{trans}_{jkm})$ was calculated assuming that transport conditions do not affect *Culicoides* viability (worst case scenario because the survival of insects is supposed to be optimal during transport). If pest control is implemented in region j , we reduced survival probabilities depending on the efficiency of the pest control product ($\text{Prot}_{\text{vect}}$) used. This was only the case for equines coming from high-risk regions [36]. Since bovines are not consistently and systematically disinfected before transport, we assumed that no pest control was implemented for them.

The probability of establishment via an infected vector was defined as:

$$P(\text{est}V_{jkm}) = P(\text{surv}_{arrival}_{jkm}) \times b_{\text{equi}_k} \times \lambda_{VH} \times [1 - [1 - \lambda_{VH} \times P(\text{surv}_{jkm}) \times b_{\text{equi}_k} \times \lambda_{HV}]^{c_{\text{idli}_{km}}}] \quad (8)$$

where $P(\text{surv}_{arrival}_{jkm})$ is the probability of an infected vector surviving to its first infectious blood meal following its arrival in area k .

The overall national and annual median probabilities were calculated using the same procedures used to calculate the PW-host probabilities.

For a more detailed description of the probabilities and parameters used, see Additional files 2 and 3.

Input data

Because accurate registration data for horses were lacking, the ratio of bovines to equines per area k (ρ_k) was estimated by combining information from different databases. The 2010 census conducted by the French Ministry of Agriculture [39] was the source for cattle and equine abundances (horses kept in agricultural settings) for each area and the IFCE-SIRE database [40] provided additional estimates of equine abundance in each area. Because it became mandatory to identify all equines in France in 2012, this database is considered to include all of the country's equines; however, dead horses are still present in the database and, as a result, the number of equines is overestimated. Two ratios were calculated—one using each of the values of equine abundance—and ρ_k was estimated in our model as a uniform distribution that ranged from the smallest to the largest ratio calculated.

The number of bovines and equines transported to France were obtained from TRACES, the TRADE Control and Expert System, which monitors the transport of animals and products of animal origin both into and within the EU [41]. In our analysis, we only included animals whose final destination was France.

Vector abundance was estimated using data from the national surveillance system implemented in France from 2009 to 2012—approximately 160 locations were surveyed to follow the activity of *Culicoides* populations [42]. The number of competent vectors feeding on a given equine in area k during month m (C_{km}) was modeled using a truncated normal distribution; μ was the average monthly number of *Culicoides* collected per overnight trap (*Culicoides imicola* and members of the *Obsoletus* complex), σ was the standard deviation, and the minimum and maximum values observed were the lower and upper bounds of the distribution, respectively. Similar parameters were used in modeling efforts by de

Koeijer et al. [43]. The average monthly temperature during month m in area k (T_{km}) was modeled using a truncated normal distribution; μ was the average temperature of each month for each year (based on daily average temperatures obtained from MARS-Agri4cast), and σ was the standard deviation, and the 1st and 99th percentile values were the lower and upper bounds of the distribution, respectively.

Analyses were performed for the three consecutive years included in the study: 2010, 2011, and 2012.

Calculations

Model calculations were performed in Microsoft Office Excel 2010 and @Risk 6.1 [44]; 10,000 iterations were run. The sensitivity analysis tool in @Risk was used to evaluate the impact of stochasticity and uncertainty in the input parameters on model results. The correlation between the values of the input parameters and the pathway-specific probabilities of introduction were calculated (Spearman's rank correlation coefficients).

The sensitivity of the model to the values of the input parameters should have been very similar across all regions and months because we used the same model and input parameter estimates, except in the case of the bovine-to-equine ratio, the temperature data, and vector abundance. Indeed, the values of these three parameters varied across regions and months (i.e., in a given month, the vector abundance could vary greatly in one region and little in another). Larger amounts of variation could have a greater impact on the model than lesser amounts of variation. The reasoning is the same for the bovine-to-equine ratio, which also varied across regions. When determining the overall probability of introduction, we thus chose to focus our sensitivity analysis on the region-time period combinations associated with the highest levels of risk and/or uncertainty.

Results

Data on equine and bovine imports

TRACES data for 2010-2012 show that, on average, 1,300 equines arrived every year in France from non-EU countries, including about forty donkeys (and no zebras). Most of these animals (close to 80 %) passed through another EU country before arriving in France. Imports from high-risk regions represented an average of 1.6 % of the total imports; imports from low- and very-low-risk regions occurred at similar levels: 45.6 % and 52.4 %, respectively. By law, bovines cannot be imported from non-EU countries.

The trading of registered horses within the EU is not required to be reported to TRACES [36]. However, it is nonetheless regularly disclosed: in the TRACES database, more than 40 % of the equines traveling from other

EU countries to France were registered horses. It is important to note that, in most of the data on the equine trade within the EU, no distinction is made between horses and donkeys. As a consequence, the TRACES database is somewhat limited in its ability to reveal equine movements within the EU. These concerns aside, according to the database, an average of 9,350 equines arrived in France every year from 2010 to 2012; 65 % came from very-low-risk regions, and 35 % came from low-risk regions. In the case of bovines, all movements are registered in the TRACES database. An average of 145,500 bovines arrived in France every year from 2010 to 2012; 61 % came from very-low-risk regions, and 39 % came from low-risk regions.

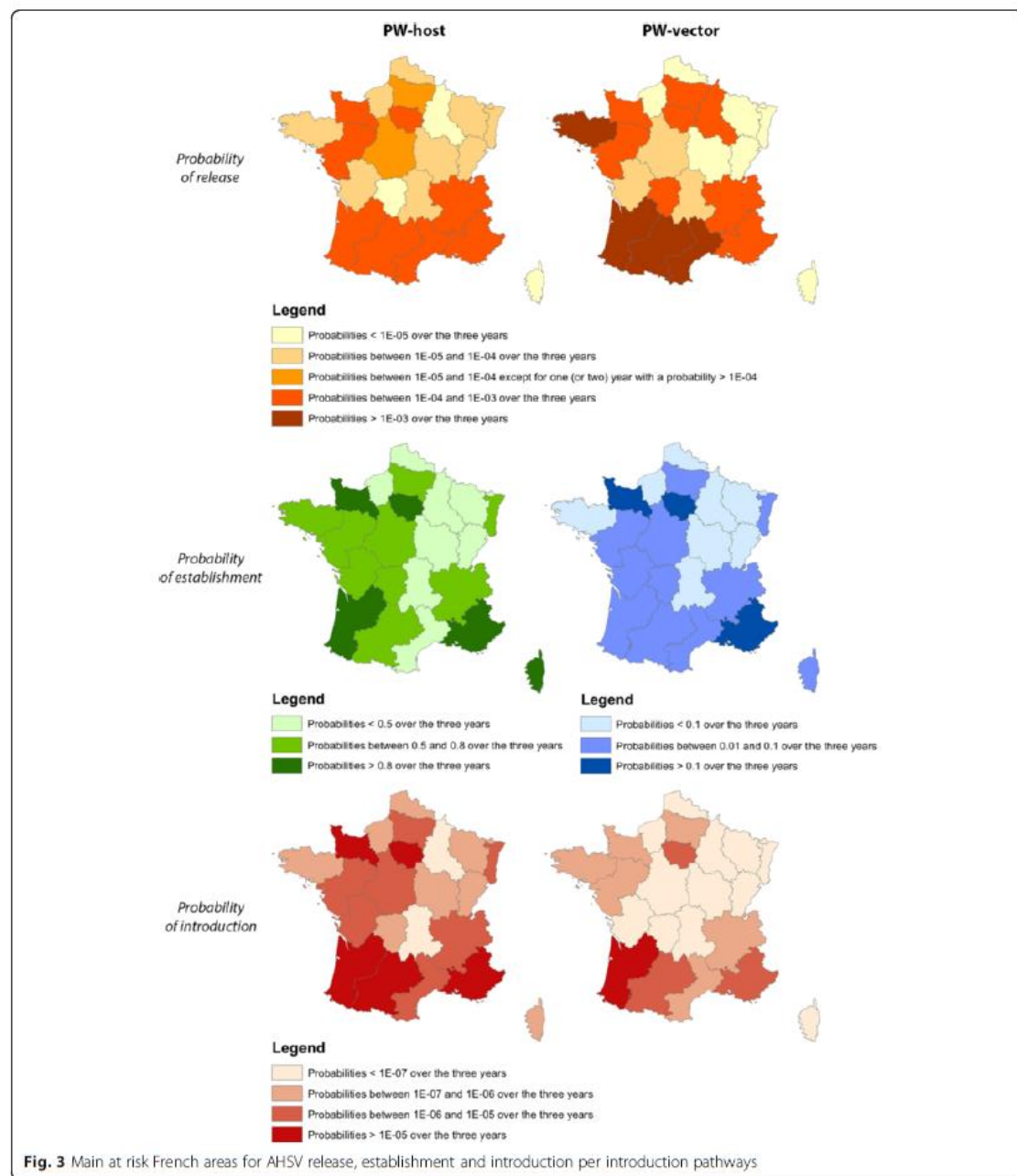
Probability of release

The probability of release is defined as the probability of an infectious equine or vector being released in a given area. The overall annual median probability of release in France was 3×10^{-3} for an infectious host (PW-host) and ranged from 1.4×10^{-2} to 3.6×10^{-2} for an infectious vector (PW-vector). Seasonal variation mostly resulted from the fact that the risk of release is negligible during the first half of the year, when low- and very-low-risk regions are considered to be unlikely to experience AHS outbreaks and equine imports from high-risk regions are very rare. From July to December across all years, the probability of release remained relatively constant; the monthly median probability that an infectious host would be released (PW-host) varied from 2.6×10^{-4} to 9.5×10^{-4} , and the monthly median probability that an infectious vector (PW-vector) would be released ranged from 1.1×10^{-3} to 6.9×10^{-3} . An exceptionally high peak was observed in July 2011 due to arrival of several horses from a high-risk country.

Areas varied greatly in their median release probabilities due to differences in the type and number of imports, but the annual probability of release for a given area was similar over time. As a result, for a given pathway of AHSV introduction, the areas most at risk remained the same (see Fig. 3).

Probability of establishment

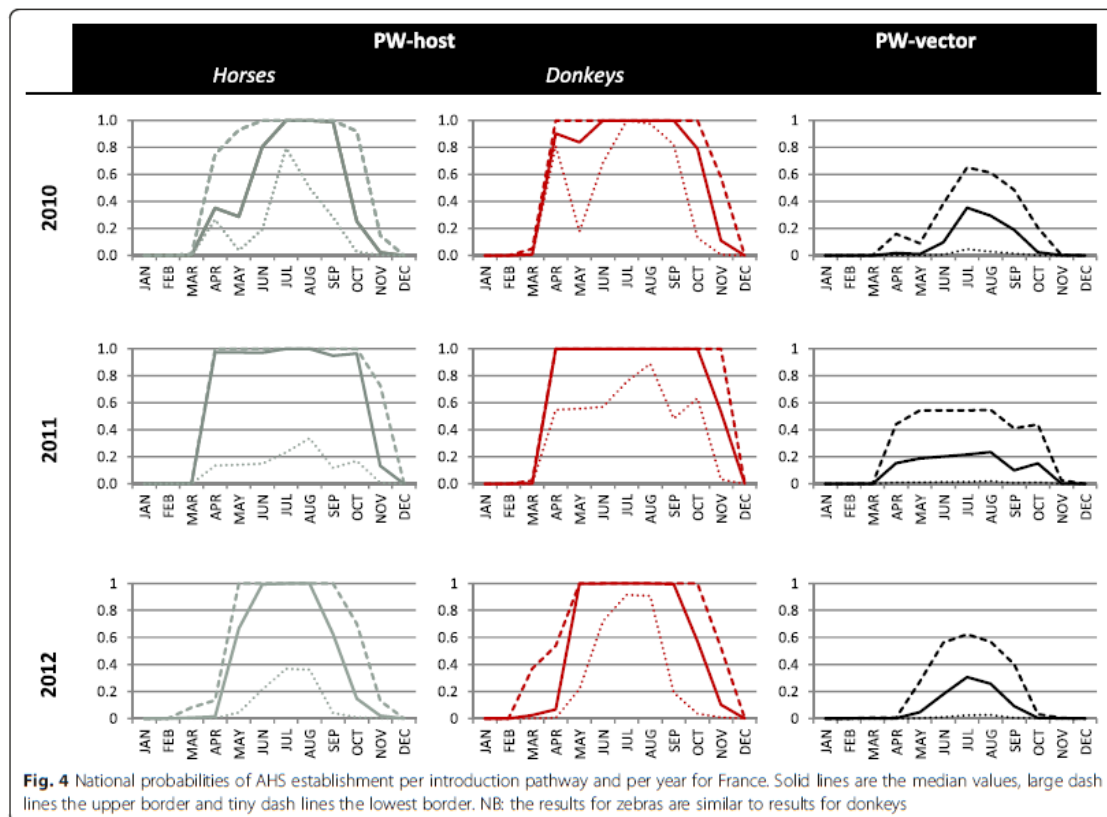
We determined the probability of establishment for each area of France, which was the probability that at least one local host would be infected by a local vector after the release of a single infectious host or vector. The probability of establishment varied as a function of temperature, vector abundance, the length of the equine host's viremic period, and the bovine-to-equine ratio in arrival area k . In the case of the infectious vector, $P(\text{est}V_{ijk})$ also depended on the life span of the specific *Culicoides* being transported. The risk of establishment



was highest from May to October and peaked between June and August (Fig. 4). Temporal and regional differences were observed—owing to variation in temperature and relative host abundance—but some areas clearly faced greater risks than others (Fig. 3).

Overall risk assessment

The probability of introduction was obtained by combining the probability of release and the probability of establishment. The median annual risk of introduction due to an infectious host was almost constant across



time (approximately 5×10^{-4}). The median annual risk of introduction due to an infectious vector varied from 4×10^{-5} to 6×10^{-5} . These figures mean that, currently, the annual risk that an infectious host will introduce AHSV into France is approximately 1 in 2,000; there is a 1 in 16,666 to 25,000 chance that AHSV introduction will be caused by an infectious vector. At the national scale, the monthly probability of introduction was similar over time, but the level of uncertainty was large (Fig. 5). The probability of introduction was the highest in the summer; it peaked in July for both pathways of introduction and in all three years. From November to June, the probability of introduction was nil, except when animals were imported from high-risk regions to the warmest areas of France (e.g., Languedoc Roussillon in March 2012). When animals were imported to colder areas, the probability of establishment was zero, making the probability of introduction zero (e.g., Basse Normandie in March 2012). Introduction risk varied greatly across space and time (see, for example, year 2012 in Fig. 6 and Additional files 4 and 5) but, over the three years examined, some areas consistently had a higher probability of

introduction (see Fig. 3). If it is assumed that an average of one midge accompanies each large animal being transported, both pathways can be combined to yield a single probability of introduction (Fig. 7), to which infectious hosts appear to be the main contributors.

The average contribution of each region of origin to introduction risk is shown in Table 1. Intra-EU trade contributes most to the risk of AHSV introduction via the infectious host pathway; low-risk EU countries are the largest contributors even though they are responsible for a lower volume of imports compared with very-low-risk EU countries. This pattern is explained by the high number of equines traded within the EU and by the fact that regulations governing intra-EU trade are less strict. No zebras were brought into France during the period we studied, and donkeys represented only 0.3 % of recorded equine imports. Their average relative contribution to AHSV introduction risk was 1.2 %. Animals imported from high-risk regions account for only 0.02 % of the large livestock arriving in France, and their average contribution to AHSV introduction risk via infectious vectors was 1.5 %. As a result, imports from

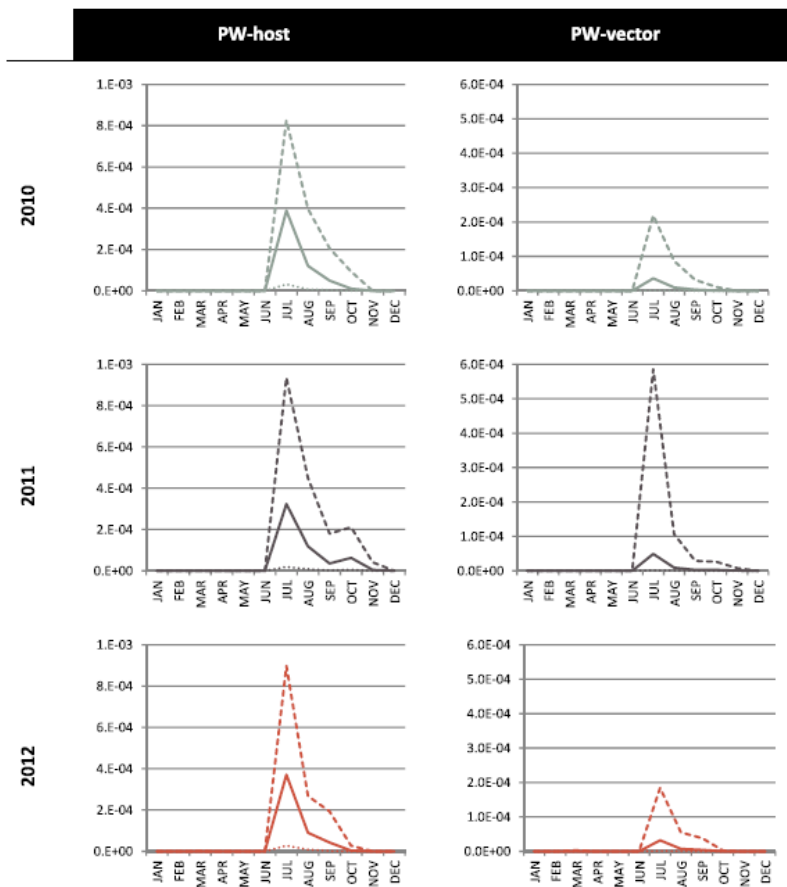


Fig. 5 National probabilities of AHSV introduction per year and per introduction pathway for France. Solid lines are the median values, large dash lines the upper border and tiny dash lines the lowest border

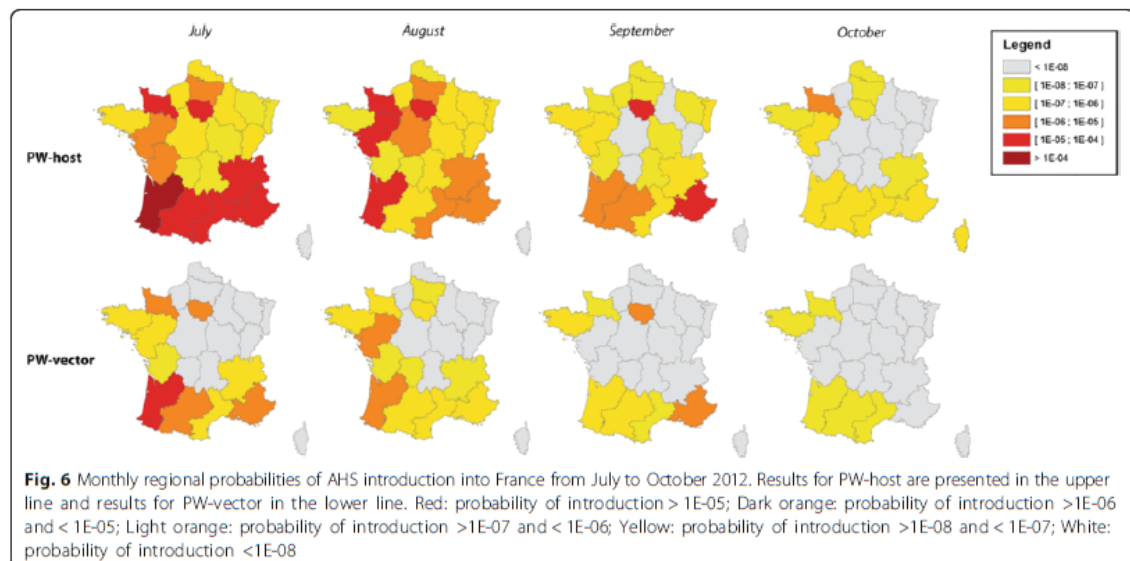
low-risk regions made the largest contribution to AHSV introduction risk via infectious vectors. Furthermore, cattle cannot be imported from outside the EU and, consequently, most of the transport of large livestock takes place within the EU (99.3 %). Given this fact and the fact that regulations regarding vector control are identical in all low-risk regions (EU and non-EU countries alike), trade of large livestock within the EU is the main contributor when it comes to the risk of AHSV being introduced by an infectious vector.

Sensitivity analysis

A sensitivity analysis was performed for the two areas identified in Fig. 7 as being at risk for AHSV introduction via the two introduction pathways: Ile de France and Provence. The level of uncertainty surrounding the risk of introduction was rather constant for both areas, with one exception: the level of uncertainty was far

higher than average in Provence in October 2012 as a result of major variation in local temperatures. The results of the analysis are summarized in Figs. 8 and 9, respectively, for the infectious host pathway and the infectious vector pathway for July (higher risk month) and October (late summer; characterized by lower risk and large uncertainty for one area) in 2012.

As expected, the values of the input parameters had a constant impact on the model's results over time and space, with only few exceptions. Furthermore, the most important input parameters mainly encompassed variability due to stochasticity (7 out of 10 parameters for PW-host and 6 out of 8 for PW-vector) and, to a lesser extent, uncertainty. Nevertheless, compared to its effects in other areas and months, average monthly temperature had a greater impact on the results for Provence in October for both pathways of introduction. This pattern was due to the fact that temperatures varied greatly in



this area during this month. This result explains the high level of uncertainty associated with the probability of introduction in this area in this month and highlights the large influence of temperature on the model's results.

Discussion

The model revealed that the annual risk of AHSV being introduced to France was very low and relatively constant for the pathways and years examined. The median value for the introduction risk via imported infectious equines (PW-host) was 0.0005; for infectious vectors (PW-vector), median introduction risk varied from 4×10^{-5} to 6×10^{-5} across years, assuming that one *Culicoides* midge arrived with each imported animal. The PW-host estimate was very similar to that obtained by de Vos and colleagues when they assessed the risk of AHSV being introduced to the Netherlands [24]. The latter study did not distinguish between EU countries and non-EU countries and only took into account competition horses; however, it did include horses traveling through the Netherlands to reach other countries. The PW-vector estimate was highly dependent on our assumption that only one *Culicoides* midge was associated with each imported animal. Indeed, the higher the number of associated *Culicoides*, the higher the probability of introduction. If it is the case that, on average, one *Culicoides* is associated with each imported animal, then the risk tied to this pathway of introduction is ten times lower than that tied to infectious hosts. As a result, AHS-free countries only face significant introduction risks if the number of *Culicoides* being transported is large as

that required for BTV, as shown by Napp et al. for Spain [19]. On the one hand, our assumption of one *Culicoides* per animal could be overly pessimistic because midges could exit the transport vehicle after feeding and thus not reach the animal's final destination. On the other hand, it could be an overly optimistic assumption because large numbers of *Culicoides* may be found on large animals. Because data are lacking on the number of *Culicoides* being transported with mammalian hosts, it is difficult to determine how each pathway of introduction contributes to overall introduction risk.

Our study indicates AHSV establishment in France may be favored from May to October. This finding is consistent with the results obtained by Lo Iacono et al. for the UK [45], by Martinez-Lopez et al. for Spain [46] and by de Vos et al. for the Netherlands [24]. In France, the favorable period for AHSV establishment is longer than in the UK (June to September) and the Netherlands (June to August) but shorter than in Spain (April to December), which is a logical consequence of climatic differences. Such differences should be taken into account when AHSV introduction within Europe is addressed at a larger scale.

Major differences were found among French regions and between introduction pathways. As expected, the coldest regions with the smallest equine populations had the lowest risk of AHSV introduction (e.g., Centre and Auvergne). In contrast, the warmest regions were most at risk for AHSV introduction (e.g., Aquitaine and Midi-Pyrénées), as were colder regions with larger equine populations (e.g., Basse Normandie and Ile de France). Warmer regions faced higher levels of risk mostly

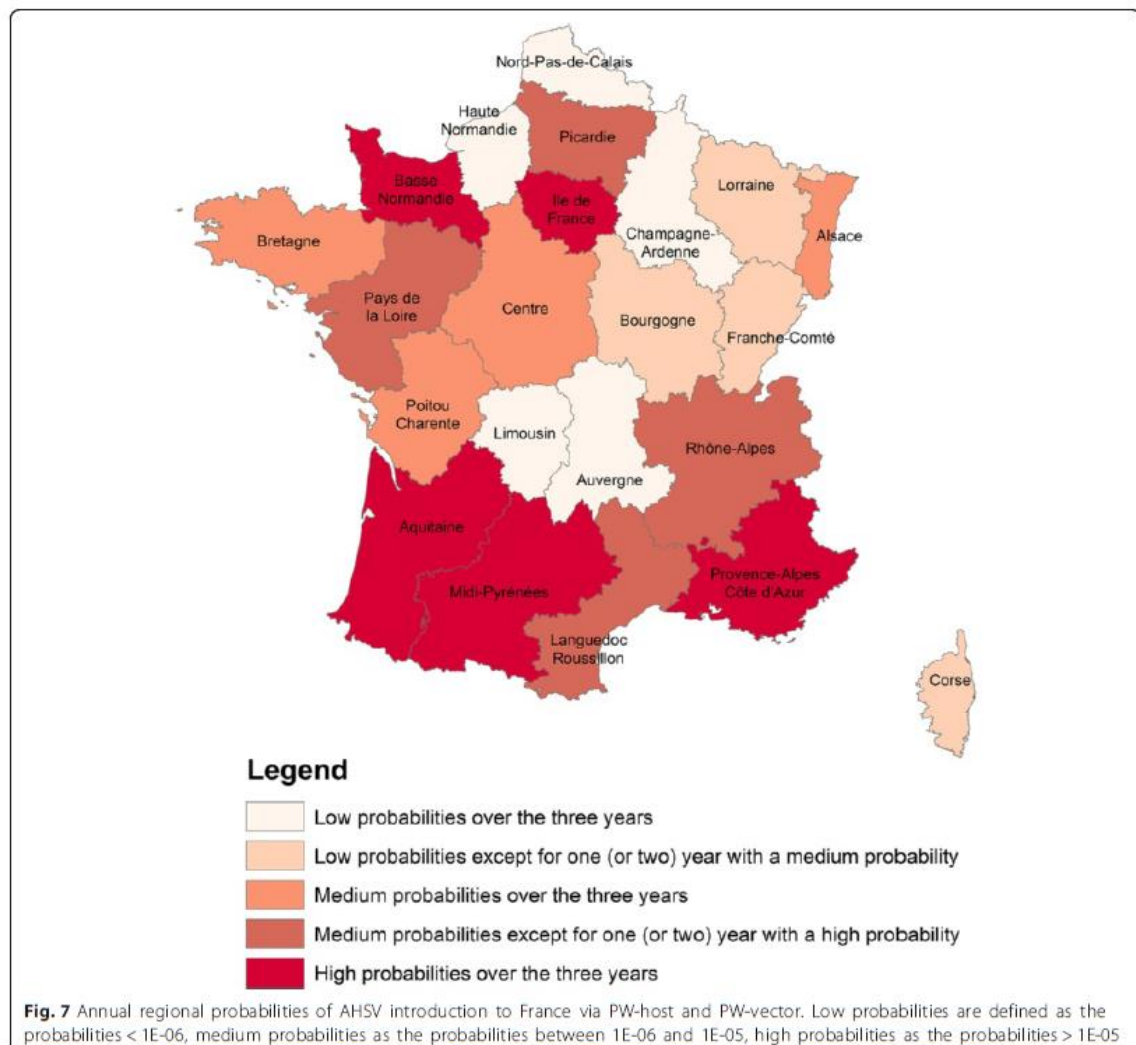
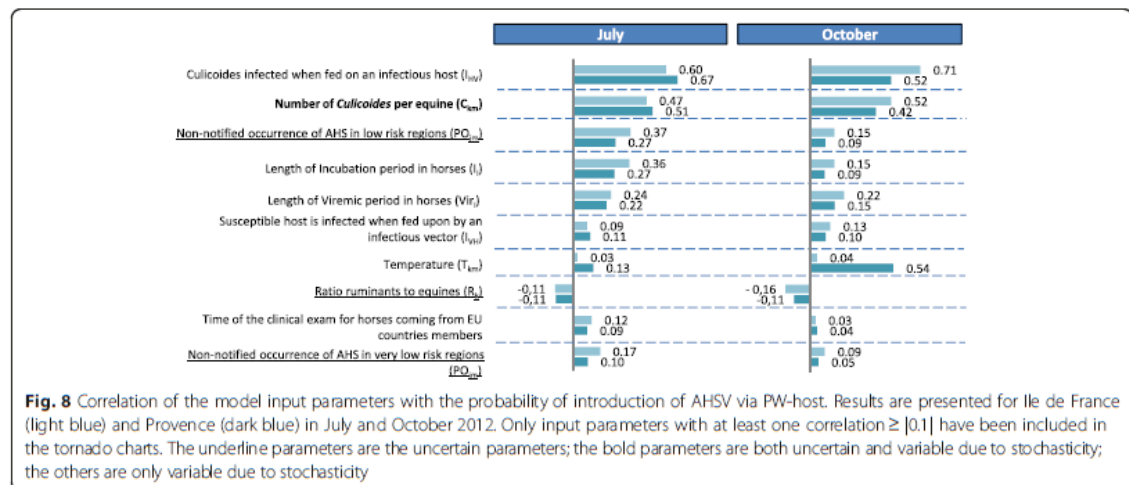


Table 1 Average contribution (%) of departure regions to the AHSV introduction risk for France. Results are presented for both pathways and compared to the total number of imports to France which are, for PW-host, the equine imports and, for PW-vector, the large animals imports (equine and bovine)

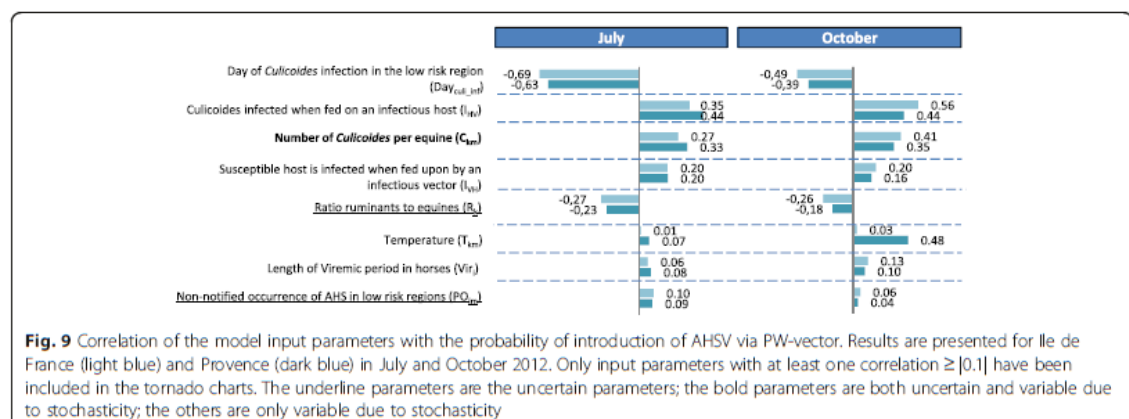
Exporting region		High risk	Low risk		Very low risk		TOTAL	
			non-EU member	EU member	non-EU member	EU member	non-EU member	EU member
PW-host	Risk	0.82	3.3	63.2	4.3	28	7.6	91.2
	Import	0.13	3.75	32.6	5.1	58.4	8.81	91
PW-vector	Risk	1.5	98.5		/		/	/
	Import	0.02	99.98		/		/	/



because of their favorable climatic conditions, while larger equine populations increased risk in colder regions. Nevertheless, if spatial differences were mainly determined by the probability of establishment, import-related variation (number and species imported) also played an important role, as seen in the *Corse* region: even if climatic conditions are favorable, the probability of introduction will be low if there are very few imported animals. These results emphasize that it is important to analyze spatiotemporal variation in risk when developing efficient surveillance systems and optimizing control measures. For instance, if there is a high risk that a pathogen will be introduced by an infectious vector, insecticides should be applied before animals are imported. In contrast, if there is a high risk that a pathogen will be introduced by an infectious host, quarantine measures should be more stringent and/or extra tests

should be performed on horses coming from low-risk regions.

We found that seasonal variation in temperature can have a large impact on the risk of introduction because it exerts a strong influence over vector abundance and biology, which, taken together, determine a vector's capacity to transmit AHSV [3, 47]. The risk of AHSV introduction was higher during periods characterized by higher than average temperatures. This finding concurs with results from work examining the introduction of BTV-8 to northwestern Europe: the extreme temperatures during July 2006 may have contributed to its widespread diffusion [48]. Therefore, rare, extreme climatic events and, more generally, global warming should have a large influence on the probability of AHSV establishment, as has been shown for BTV [49, 50]. Given the progression of global warming, risk assessments



should be regularly updated to account for climatic changes.

Our study reveals that complete and accurate data on the movements and distribution of the EU's equine population are not available: it is hard to trace horses within the EU. The introduction of mandatory horse passports in 2008 improved the situation but, apart from rare exceptions [46], it is still difficult to follow the EU's equine population [51] and assessing the population's distribution and fluxes remains a challenge (see the UK [52]). for an example). This is a major concern given that the distribution of the equine population had an important impact on our results (see the ratio of bovines to equines in Figs. 8 and 9); moreover, the number of equines being imported is obviously highly correlated with introduction probabilities. Furthermore, several equine viruses are zoonotic (e.g., eastern and western equine encephalomyelitis viruses, Venezuelan equine encephalitis virus, and West Nile virus), and the risk that they will be introduced to and spread within the EU is definitely not negligible [22]. It is, in fact, currently increasing; indeed, West Nile virus has already become endemic in some regions [53]. Improving the traceability of horses within the EU would thus be advantageous when it comes to better assessing the risk posed by AHSV and other zoonotic diseases.

The risk assessment model described in this paper addresses the risk of AHSV being introduced to France by two pathways of introduction considered to be of importance [23]. However, other pathways may also substantially contribute to introduction risk. Several studies have highlighted that the wind may efficiently transport *Culicoides* over long distances, both across sea [25, 54] and land [55]; it might have been involved in the spread of BTV in Europe [25–27]. Wind-mediated dispersal of infected vectors might also have resulted in AHSV being introduced to Cape Verde Island in 1944, Cyprus in 1960, the Middle East in 1960 and Spain in 1966 [56]. An extensive assessment of the role played by the wind in spreading *Culicoides* midges and *Culicoides*-borne infections across the Mediterranean Basin would elucidate the importance of this pathway for AHS introduction. However, the wind-mediated dispersal of AHSV is most likely to occur in the south of France close to low-risk regions, which is also where AHSV introduction is most likely to occur via the pathways examined in this study. Therefore, we think that this study accurately identifies the regions of France that face the greatest risk of AHSV introduction. Furthermore, the risk that AHSV will be introduced via wind-borne infectious *Culicoides* cannot be mitigated by direct preventive measures, such as importation restrictions. Instead, to be effective, control measures would have to influence the probability of establishment;

for instance, insecticides could be used to protect local hosts against wind-borne vectors.

Our model provides a basis for creating a risk-based surveillance system in France that focuses on the regions and time periods associated with higher levels of AHSV introduction risk. The model could also be used for assessing risk and establishing surveillance procedures in other European countries. This application is especially important because our study has revealed that European countries make the largest contribution to France's AHSV introduction risk (e.g., PW-host: 91.2 %). Indeed, if an infection occurs in one European country but is not detected, then it can easily spread to other European countries because there is little verification and tracking of equine movements within the EU. By implementing a risk-based surveillance strategy in each country in the EU, infections would have a higher probability of being detected early on; as a consequence, the contribution of fellow EU countries to introduction risk would decline (see the probabilities of non-notified AHS occurrence in low- and very-low-risk regions in Figs. 8 and 9). By reinforcing the tracking of equine movements within the EU, infection would also have less chance to disseminate and the policy implications of an AHS introduction will be more limited. Such strategies could be a means of minimizing the risk and impact of an AHSV outbreak for the entire European equine industry.

Conclusion

We have developed a quantitative risk assessment model to estimate the risk of AHSV being introduced to France via the importation of infectious equines and infectious *Culicoides* midges associated with large livestock. The risk that AHS will be introduced to France is very low; however, risk varies tremendously among the different regions of the country due to variation in temperature and equine population size. The regions most at risk are those with the warmest climates as well as those that are colder but that harbor larger equine populations. Introduction risk is greatest from July to October and peaks in July. Despite the low probability that AHSV is present in the EU, intra-EU trade of equines contributes most to the risk of AHSV introduction to France because it is responsible for a large number of horse movements. Spatiotemporal differences need to be addressed when assessing the risk that AHSV will be introduced to a given location and when developing and implementing risk-based surveillance procedures. The methods and results of this study may help guide surveillance programs and other risk-reduction measures aimed at preventing the introduction of AHSV or minimizing its potential impact once it has been introduced, both in France and in other European countries.

Additional files

Additional file 1: Model calculation for PW-host. Details of calculation regarding the AHSV introduction via the import of an infectious host.

Additional file 2: Model parameters. Description of all parameters used in the model calculation for PW-host and PW-vector.

Additional file 3: Model calculation for PW-vector. Details of calculation regarding the AHSV introduction via the import of an infectious vector.

Additional file 4: Spatiotemporal risk of AHSV introduction via PW-host to France from 2010 to 2012.

Additional file 5: Spatiotemporal risk of AHSV introduction via PW-vector to France from 2010 to 2012.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

CF participated in the design of the study, performed the statistical analysis and drafted the manuscript; AL participated in data acquisition and interpretation and contributed to the writing of the manuscript; PH participated in acquisition of the data and contributed to the improvement of the manuscript; TB participated in acquisition of the data and contributed to the improvement of the manuscript; CDV participated in the interpretation of data and contributed to the writing of the manuscript; EF participated in the interpretation of data and contributed to the writing of the manuscript; ADK participated in the design and coordination of the study and contributed to the writing of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

Entomological data were collected by the National entomological surveillance system that is funded by the French Ministry in charge of agriculture and implemented by CIRAD, the *Entente interdépartementale pour la démosculation du littoral méditerranéen* (EID-Med), and the *Institut de parasitologie et de pathologie tropicale de Strasbourg* (IPPTS) with the support of the *Directions départementales de la protection des populations*. Authors want to thank all persons involved in this network and especially those in charge of insect identification: Claire Garros, Xavier Allène, Ignace Rakotariyony, Jonathan Lhoir (Cirad), Jean-Claude Delécolle, Delphine Delécolle, Bruno Mathieu (IPPTS), Marie-Laure Setier-Rio, Bethsabée Scheid and Roger Venail (EID-Med). Temperature data were kindly provided by MARS-Agri4cast at the Institute for Environment and Sustainability, European Commission. Data on trade in live animals were provided by the *Bureau de l'identification et du contrôle des mouvements des animaux* (BICMA) and the *service d'inspection vétérinaire et phytosanitaire aux frontières* (SIVP) at the General Direction for Food (DGAI) of the French Ministry of Agriculture. The authors want to thank all these institutes for providing these data.

This research was carried within the VICE project funded under the FP7 EMIDA ERA-NET initiative. National funding within ERA-Net was provided by the Dutch Ministry of Economic Affairs (Project nr. BO-20-009-009). This work was also partially financially supported by IFCE (Institut Français du Cheval et de l'Équitation) and Vetagro Sup (National Veterinary School of Lyon).

Author details

¹INRA UR346 Animal Epidemiology, Vetagrosup, F-69280 Marcy l'Etoile, France. ²INRA UR346 Animal Epidemiology et Département Hippique, VetAgroSup, F-69280 Marcy l'Etoile, France. ³ANSES, Direction scientifique des laboratoires – unité Survepi, 94700 Maisons-Alfort, France. ⁴CIRAD, UMR CMAEE, F-34398 Montpellier, France ; INRA, UMR1309 CMAEE, F-34398 Montpellier, France. ⁵Central Veterinary Institute, part of Wageningen UR, PO Box 658200 AB Lelystad, The Netherlands.

Received: 1 September 2014 Accepted: 12 May 2015

Published online: 04 June 2015

References

- Mellor PS. African horse sickness: transmission and epidemiology. *Vet Res*. 1993;2:199–212.

- Du Toit RM. The transmission of bluetongue and horse sickness by *Culicoides*. *Onderstepoort J Vet Sci Anim Ind*. 1944;19:7–16.
- Wilson A, Mellor PS, Szmaragd C, Mertens PPC. Adaptive strategies of African horse sickness virus to facilitate vector transmission. *Vet Res*. 2009;40(2):16.
- Mellor PS, Hamblin C. African horse sickness. *Vet Res*. 2004;35:445–66.
- Barnard BJ. Circulation of African horsesickness virus in zebra (*Equus burchelli*) in the Kruger National Park, South Africa, as measured by the prevalence of type specific antibodies. *Onderstepoort J Vet Res*. 1993;60:111–7.
- MacLachlan NJ, Guthrie AJ. Re-emergence of bluetongue, African horse sickness, and other Orbivirus diseases. *Vet Res*. 2010;41(6):35.
- Rodríguez M, Hooghuis H, Castaño M. African horse sickness in Spain. *Vet Microbiol*. 1992;33:129–42.
- Zientara S, Sánchez-Vizcaino JM. Control of bluetongue in Europe. *Vet Microbiol*. 2013;166:33–7.
- Doceul V, Lara E, Sailleau C, Belbis G, Richardson J, Breard E, et al. Epidemiology, molecular virology and diagnostics of Schmallenberg virus, an emerging orthobunyavirus in Europe. *Vet Res*. 2013;44:31.
- Carpenter S. Editorial: Bluetongue, Schmallenberg virus... African horse sickness? *Vet Rec*. 2014;174:299–300.
- Lefèvre P-C, Blancou J, Chermette R, Uilenberg G. *Infectious and Parasitic Diseases of Livestock, Volume 1*. Paris: Lavoisier; 2010:2010.
- Venail R, Balenghien T, Guis H, Tran A, Setier-Rio M-L, Delécolle J-C, Mathieu B, Cêtre-Sossah C, Martinez D, Langui J, Baldet T, Garros C. Assessing Diversity and Abundance of Vector Populations at a National Scale: Example of *Culicoides* Surveillance in France After Bluetongue Virus Emergence. In *Arthropods Vectors Emerg Dis. Volume 3, Chapter 4*. Springer-Verlag, Berlin Heidelberg; 2012:77–102. [Parasitology Research Monographs]
- Lebrun J. *Les enjeux et les perspectives de la filière équine en France. Rapport public*. Conseil économique, social et environnemental. Section de l'agriculture et de l'alimentation; 2010. Available at: <http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/104000387.pdf>
- Allison K, Taylor N, Upton M, Wilmshire T. *African Horse Sickness, Impact on the UK Horse Industry*. Veterinary Epidemiology and Economics Research Unit, University of Reading; 2009. Available at: <http://www.veeru.reading.ac.uk/documents/AHS%20-%20HT%20-%20Report%20final%20301109.pdf> (accessed 31.05.13)
- Thompson GM, Jess S, Murchie AK. A review of African horse sickness and its implications for Ireland. *Ir Vet J*. 2012;65:9.
- Mourits MCM, Saatkamp HW. *Kostenberekening van een uitbraak met Afrikaanse paardenpest in Nederland*. Bedrijfs-economische Wageningen Universiteit; 2010. Available at: <http://edepot.wur.nl/168958> (accessed 10.06.14 in Dutch)
- Sutmoller P, Wrathall AE. A quantitative assessment of the risk of transmission of foot-and-mouth disease, bluetongue and vesicular stomatitis by embryo transfer in cattle. *Prev Vet Med*. 1997;32:111–32.
- Hoar B, Carpenter T, Singer R, Gardner I. Probability of introduction of exotic strains of bluetongue virus into the US and into California through importation of infected cattle. *Prev Vet Med*. 2004;66:79–91.
- Napp S, García-Bocanegra I, Pagès N, Allepuz A, Alba A, Casal J. Assessment of the risk of a bluetongue outbreak in Europe caused by *Culicoides* midges introduced through intracontinental transport and trade networks. *Med Vet Entomol*. 2012;27:19–28.
- Kilpatrick AM, Gluzberg Y, Burgett J, Daszak P. Quantitative risk assessment of the pathways by which west Nile virus could reach Hawaii. *EcoHealth*. 2004;1:205–9.
- Brown EBE, Adkin A, Fooks AR, Stephenson B, Medlock JM, Snary EL. Assessing the risks of west Nile virus-infected mosquitoes from transatlantic aircraft: implications for disease emergence in the united kingdom. *Vector-Borne Zoonotic Dis*. 2012;12:310–20.
- Durand B, Lecollinet S, Beck C, Martínez-López B, Balenghien T, Chevalier V. Identification of hotspots in the European union for the introduction of four zoonotic arboviruses by live animal trade. *PLoS One*. 2013;8:e70000.
- Sabirovic M, Lopez M, Patel K, Kingston A, Hall S. *African Horse Sickness: Potential Risk Factors and the Likelihood for the Introduction of the Disease to the United Kingdom*. Food and Farming Group, Department of Environment Food and Rural Affairs; 2008. Available at: <http://webarchive.nationalarchives.gov.uk/20130402151656/http://archivedefra.gov.uk/>

- foodfarm/farmanimal/diseases/monitoring/documents/ahs_uk081106.pdf (accessed 12.11.12)
24. De Vos CJ, Hoek CA, Nodelijk G. Risk of introducing African horse sickness virus into the Netherlands by international equine movements. *Prev Vet Med.* 2012;106:108–22.
 25. Burgin LE, Gloster J, Sanders C, Mellor PS, Gubbins S, Carpenter S. Investigating incursions of bluetongue virus using a model of long-distance culicoides biting midge dispersal. *Transbound Emerg Dis.* 2012;60:263–72.
 26. Ågren ECC, Burgin L, Lewerin SS, Gloster J, Elvander M. Possible means of introduction of bluetongue virus serotype 8 (BTV-8) to Sweden in August 2008: comparison of results from two models for atmospheric transport of the Culicoides vector. *Vet Rec.* 2010;167:484–8.
 27. Ducheyne E, De Deken R, Bécu S, Codina B, Nomikou K, Mangana-Vougiaki O, et al. Quantifying the wind dispersal of Culicoides species in Greece and Bulgaria. *Geospatial Health.* 2007;1:177–89.
 28. Carpenter S, Wilson A, Mellor PS. Culicoides and the emergence of bluetongue virus in northern Europe. *Trends Microbiol.* 2009;17:172–8.
 29. Mintiens K, Méroc E, Mellor PS, Staubach C, Gerbier G, Elbers ARW, et al. Possible routes of introduction of bluetongue virus serotype 8 into the epicentre of the 2006 epidemic in north-western Europe. *Prev Vet Med.* 2008;87:131–44.
 30. De La Rocque S, Balenghien T, Halos L, Dietze K, Claes F, Ferrari G, et al. A review of trends in the distribution of vector-borne diseases: Is international trade contributing to their spread? *Rev Sci Tech Int Off Epizoot.* 2011;30:119–30.
 31. De Vos CJ, Hoek M, Fischer E, De Koeijer A, Bremmer J. *Risk Assessment Framework for Emerging Vector-Borne Livestock Diseases.* Netherlands: Central Veterinary Institute of Wageningen UR; 2011. Available at: <https://www.wageningenur.nl/vervPublication-details.htm?publicationId=publication-way-343232303431> (accessed 05.01.13)
 32. Ninio C, Augot D, Delecote J-C, Dufour B, Depaquit J. Contribution to the knowledge of culicoides (diptera: ceratopogonidae) host preferences in France. *Parasitol Res.* 2011;108:657–63.
 33. Viennet E, Garros C, Gardès L, Rakotoarivony I, Allène X, Lancelot R, et al. Host preferences of Palaearctic Culicoides biting midges: implications for transmission of orbiviruses. *Med Vet Entomol.* 2013;27:255–66.
 34. Gratz NG, Steffen R, Cockedge W. Why aircraft disinsection? *Bull World Health Organ.* 2000;78:995–1004.
 35. Mellor PS, Boorman J, Baylis M. Culicoides biting midges: their role as arbovirus vectors. *Annu Rev Entomol.* 2000;45:307–40.
 36. CEC (Commission of the European Communities). Council Directive 2009/156/EC of 30 November 2009 on Animal Health Conditions Governing the Movement and Importation from Third Countries of Equidae. *Off J Eur Union.* 2010 L192/1
 37. Commission Decision (EC) 2004/211 of 6 January 2004 establishing the List of Third Countries and Parts of Territory Thereof from Which Member States Authorise Imports of Live Equidae and Semen, Ova and Embryos of the Equine Species, and Amending Decisions 93/195/EEC and 94/63/EC [2004] OJ L73/1
 38. Stevenson P, Formosinho J. *Long Distance Animal Transport in Europe: A Cruel and Unnecessary Trade.* Compassion in World Farming; 2008. Available at: <https://www.ciwf.org.uk/media/3818249/transport-in-europe-report.pdf> (accessed 06.06.14)
 39. Agreste, 2010. Recensement Agricole 2010. Available at: <http://agreste.agriculture.gouv.fr/recensement-agricole-2010/resultats-donnees-chiffres/> (accessed 11.12.12 in French)
 40. REFERENCES, 2011. Les effectifs d'équidés présents en France. Available at: <http://www.haras-nationaux.fr/information/accueil-equipedia/filiere-economie/chiffres-des-sur-les-entreprises-ressources-et-territoire/les-effectifs-dequides-en-france.html?type=98> (accessed 11.12.12 in French)
 41. Commission Decision (EC) 2003/623 of 19 August 2003 concerning the Development of an Integrated Computerised Veterinary System Known as Traces [2003] OJ L216/58
 42. Balenghien T, Delécolle J-C, Setier-Rio M-L, Rakotoarivony I, Allène X, Venail R, et al. Vecteurs du virus de la fièvre catarrhale ovine: suivi des populations de culicoides en 2011 en France. *Bull Épidémiologique Santé Anim Aliment.* 2012;54:35–40.
 43. De Koeijer A, Hartemink N, Boender GJ, Elbers A, Heesterbeek H. *Epidemiological Analysis of the 2006 Bluetongue Virus Serotype 8 Epidemic in North-Western Europe: A Risk Map for Epidemic Potential in the Netherlands.* Animal Science group. Wageningen UR; 2007. Available at: <http://www.efsa.europa.eu/fr/efsajournal/doc/34r.pdf> (accessed 06.06.14)
 44. Palisade Corporation. Guide to using @Risk. Risk Analysis and Simulation Add-In for Microsoft Excel, Version 6.1. Newfield, NY, USA; Palisade Corporation; 2013
 45. Lo Iacono G, Robin CA, Newton JR, Gubbins S, Wood JLN. Where are the horses? With the sheep or cows? Uncertain host location, vector-feeding preferences and the risk of African horse sickness transmission in Great Britain. *J R Soc Interface.* 2013;10:20130194–4.
 46. Martínez-López B, Perez AM, Sánchez-Vizcaíno JM. Identifying equine premises at high risk of introduction of vector-borne diseases using geo-statistical and space-time analyses. *Prev Vet Med.* 2011;100:100–8.
 47. Carpenter S, Wilson A, Barber J, Veronesi E, Mellor P, Venter G, et al. Temperature dependence of the extrinsic incubation period of orbiviruses in Culicoides biting midges. *PLoS One.* 2011;6:e27987.
 48. Wilson AJ, Mellor PS. Bluetongue in Europe: past, present and future. *Philos Trans R Soc B Biol Sci.* 2009;364:2669–81.
 49. Purse BV, Mellor PS, Rogers DJ, Samuel AR, Mertens PPC, Baylis M. Climate change and the recent emergence of bluetongue in Europe. *Nat Rev Microbiol.* 2005;3:171–81.
 50. Guis H, Caminade C, Calvete C, Morse AP, Tran A, Baylis M. Modelling the effects of past and future climate on the risk of bluetongue emergence in Europe. *J R Soc Interface.* 2012;9:339–50.
 51. Liljenstolpe C. *Horses in Europe.* Conference EU Equus 2009: Swedish University of Agricultural Sciences; 2009. Available at: <http://www.wbfs.org/files/EU%20Equus%202009.pdf> (accessed 10.12.12)
 52. Boden LA, Parkin TD, Yates J, Mellor D, Kao RR. An online survey of horse-owners in Great Britain. *BMC Vet Res.* 2013;9:188.
 53. Calistri P, Giovannini A, Hubalek Z, Ionescu A, Monaco F, Savini G, et al. Epidemiology of West Nile in Europe and in the Mediterranean Basin. *Open Virol J.* 2010;4:29–37.
 54. Eagles D, Walker PJ, Zalucki MP, Durr PA. Modelling spatio-temporal patterns of long-distance Culicoides dispersal into northern Australia. *Prev Vet Med.* 2013;110:312–22.
 55. García-Lastra R, Leginagoikoa I, Plazaola JM, Ocabo B, Aduriz G, Nunes T, et al. Bluetongue Virus Serotype 1 Outbreak in the Basque Country (Northern Spain) 2007–2008. Data Support a Primary Vector Windborne Transport. *Plos One* 2012;7(3):e3442
 56. Sellers RF, Pedgley DE, Tucker MR. Possible spread of African horse sickness on the wind. *J Hyg (Lond).* 1977;79:279–98.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



C. COMPARISON OF DISEASES RISKS

This study aimed first to evaluate the feasibility of adapting the model developed in Chapter II.B to another disease. For that purpose, we chose equine encephalosis (EE) as this disease is similar to AHS. Then, we compared the probabilities of entry into France of both diseases in order to better understand and manage the risk associated with AHS and EE.

This part of the work resulted in two poster presentations at the Conference of the Society for Veterinary Epidemiology and Preventive Medicine (SVEPM) in Ghent (2015) and at the *Journée de l'Ecole doctorale* in Clermont-Ferrand (2015) (posters available in Appendix 7 and Appendix 8, respectively) and in one oral presentation at the final workshop of the EMIDA-VICE project (February 2015).

1. Introduction

In Europe, African horse sickness (AHS) and equine encephalosis (EE) have been recently highlighted as potential threats (MacLachlan and Guthrie 2010; Zimmerli et al. 2010). Like bluetongue and Schmallenberg viruses, both are borne by biting midges in the genus *Culicoides*, and both are non-zoonotic equid diseases caused by viruses belonging to the *Orbivirus* genus of the Reoviridae family (Dhama et al. 2014; Mellor and Hamblin 2004; Viljoen and Huismans 1989). Both viruses have more or less the same geographical distribution and are considered endemic in sub-Saharan countries, with rare outbreaks in North Africa and western Asia (MacLachlan and Guthrie 2010; Mildenberg et al. 2009; Wescott et al. 2013). Likewise, the epidemiology of these diseases are also similar, with both sharing the same vectors (Venter et al. 2002; Venter et al. 2000), although EE has a higher transmission rate (Lord et al. 2002). However, the two viruses differ in pathogenicity: AHS is one of the most devastating diseases in equids, with a mortality rate approaching 90% (Mellor and Hamblin 2004), whereas EE induces only sporadic symptoms with a correspondingly low mortality rate (Dhama et al. 2014). At first glance, the epidemiology and transmission patterns of AHS and EE are very similar; however, more detailed information on the respective probabilities of entry of the two diseases is needed.

The aim of this study was to determine the extent to which AHS and EE differ in their probabilities of entry, despite their similar patterns of epidemiology and transmission. We also identified the most appropriate measure for risk mitigation for each disease. The probabilities of entry into France of both AHS and EE were evaluated and compared for two routes of virus entry: an infectious host imported via legal trade and an infectious vector imported via the large animal trade.

2. Method

2.1. Model for risk assessment of viruses entry

The probability of entry is defined as the probability that a virus reaches an area, without consideration of later steps of transmission (OIE 2014). The stochastic spatiotemporal model of AHS introduction presented in the previous section (Chapter 2, section B)(Faverjon et al. 2015b) was used to assess the probability of AHS entry. The model was adapted by Evelyn Pamela Martinez Lopez, a Master's student from CVI, to assess the probability of EE introduction to the Netherlands and subsequently adapted by us to assess the probability of EE entry into France. Two pathways were considered: the introduction of an infectious host via legal trade and the introduction of an infectious vector through the large animal trade.

2.2. Parameters

The models for AHS and EE entry differ only in the disease parameters. The parameters not related to the diseases, such as transport time, were identical to those defined in the previous section (Chapter 2, section B) (Faverjon et al. 2015b). The disease-specific parameters were, for AHS, those used in the previous section (Chapter 2, section B) (Faverjon et al. 2015b) and, for EE, the parameters used by Evelyn Pamela Martinez Lopez. The differences between the diseases are reported in Table 3 and Table 4. All other parameters are identical in both models; in particular, there is no quarantine and no laboratory test implemented for either disease for horses traveling within the EU.

Table 3 : import procedure for African horse sickness and equine encephalosis

	African horse sickness	Equine encephalosis
High risk	Quarantine: 40 days 2 ELISA tests with minimum 21 days and maximum 30 days between them (Sensitivity: Beta(60, 4), Specificity: Beta(62, 2)) Vectors protection (efficacy Uniform(0.5,0.9)) Time to clinical inspection: day of embarkation	Quarantine: 40 days 2 ELISA tests with minimum 21 days and maximum 30 days between them (Sensitivity: 1, Specificity: 1) Vectors protection (efficacy Uniform(0.5,0.9)) Time to clinical inspection: day of embarkation <i>Israel:</i> Quarantine: 40 days No ELISA test. Vectors protection (efficacy Uniform(0.5,0.9)) Time to clinical inspection: day of embarkation
Low risk (non EU countries)	Quarantine: 40 days 2 ELISA tests with minimum 21 days and maximum 30 days between them (Sensitivity: Beta(60, 4) Specificity: Beta(62, 2)) No Vectors protection Time to clinical inspection: day of embarkation	Quarantine: 40 days No laboratory test. No vector protection Time to clinical inspection: day of embarkation
Very low risk (non EU countries)	No quarantine, No laboratory test, No vector protection Time to clinical inspection: at least 48h before embarkation	No quarantine, No laboratory test, No vector protection Time to clinical inspection: at least 48h before embarkation

Table 4 : Parameters specific of diseases and used for African horse sickness and Equine encephalosis in the model of viruses entry. The parameters used for African horse sickness come from (Faverjon et al. 2015b); The parameters used for equine encephalosis have been estimated following the same assumptions than those used in (Faverjon et al. 2015b) but adapted according to the specificities of the disease presented in (Aharonson-Raz et al. 2011) (Crafford et al. 2003)(Crafford et al. 2011)(Mildenberg et al. 2009)(Paweska and Venter 2004)(Venter et al. 1999)

		African horse sickness	Equine encephalosis
Mortality rate for hosts		Horse: 70%, Donkey: 10%, Zebra: 1%	All equidae: 5%
Length of viraemia (days)		<i>When animal died:</i> Horse: Gamma(20.25, 0.22) Donkey: 12 Zebra: 28 <i>When animal survived:</i> Horse: Gamma(29.75, 0.20) Donkey: 28 Zebra: 40	All equidae: Pert(7, 30) Mean 19
Incubation period (days)		Pert(2, 10) Mean 6	Pert(2, 6) Mean 4
Time between infection and seroconversion (days)		Uniform(10, 14)	Uniform(length of incubation, length of viraemia)
Definition of at risk area	High risk	Sub-Saharan countries	Sub-Saharan countries + Israel
	Low risk	Regions that have experienced AHS outbreaks in the past and/or where the main vector, <i>C. imicola</i> , is present	Regions that have experienced AHS outbreaks in the past and/or where the main vector, <i>C. imicola</i> , is present
	Very low risk	All the other countries	All the other countries
Rate of Culicoides infected during an outbreak	High risk	0.014	Uniform(0.0005, 0.23)
	Low risk	0.00014	Uniform(0.0005, 0.23)/100
High risk period (days)	Low risk	2*Incubation period for AHS + time till next infectious blood meal based on 18°C	70 days
	Very low risk	2*Incubation period for AHS + time till next infectious blood meal based on 12°C	70 days
Probability of disease occurrence	Low risk	Gamma(AHS High risk period in Low risk region*15, 1/(365*60)) Mean: 0.016	Gamma(7*365, 1/(365*47)) Mean: 0.14
	Very low risk	Gamma(AHS High risk period in Very Low risk region, 1/(365*61)) Mean: 0.0027	Gamma(EE High risk period in Very Low risk region, 1/(365*48)) Mean: 0.0033
Cumulative incidence	High risk	Horse: Pert(4×10^{-6} , 5.02×10^{-4} , 1×10^{-3}) Donkey: 1.2×10^{-2} Zebra: 1.6×10^{-2}	All equidae: $1 - \exp(\text{Uniform}(0.29, 0.67))$
	Low risk	2×10^{-4}	0.0951
	Very low risk	2×10^{-4}	0.0487

2.3. Input data

To calculate the probability of virus release, the numbers of bovines and equines transported to France were obtained from TRACES, the **TRAde Control and Expert System**, which monitors the transport of animals and products of animal origin both into and within the EU. In our analysis, we only included animals whose final destination was France.

2.4. What-if scenarios

To evaluate the impact of manageable parameters on the probability of entry of AHS and EE into France, we used the average annual number of animals imported into France between 2010 and 2012 to evaluate different scenarios and compare them to the default. Seven scenarios were tested: five tested the effects of changes to existing legislation, and two evaluated the consequences of changing import procedures from non-EU countries. A description of the scenarios is available in Table 5.

Table 5 : Description of scenario tested. Scenarios 1 to 5 changed the existing legislation, scenarios 6 to 7 changes imports from third countries.

Scenario 1	Quarantine period of 60 days instead of 40
Scenario 2	Quarantine period of 20 days instead of 40
Scenario 3	<p>Same regulation implemented in Low risk EU state than in low risk non-EU state</p> <ol style="list-style-type: none"> AHS: Quarantine of 40 days, 2 ELISA tests with minimum 21 days and maximum 30 days between them (Sensitivity: Beta(60, 4) Specificity: Beta(62, 2)), no Vectors protection, and clinical inspection the day of embarkation. EE: Quarantine of 40 days, no laboratory test, no vector protection, and clinical inspection the day of embarkation.
Scenario 4	No serological test implemented
Scenario 5	Vector protection implemented on all animals coming from low risk area
Scenario 6	No legal importation from high risk region
Scenario 7	No legal importation from low risk region and non-EU states

2.5. Calculations

Model calculations were performed in Microsoft Office Excel 2010 with @Risk 6.1 (Palisade Corporation 2013); 10,000 iterations were run. The sensitivity analysis tool in @Risk was used to evaluate the impact of stochasticity and uncertainty in the input parameters on model results. The correlations between the values of the input parameters and the pathway-specific probabilities of introduction were calculated (Spearman's rank correlation coefficients).

The sensitivity of the model to input parameter values could be expected as very similar across all regions and months because we used the same model and input parameter estimates in every case, with the exception of the bovine-to-equine ratio, the temperature data, and vector abundance. The values of these three parameters varied substantially across regions and months (e.g., in a given month, vector abundance could vary greatly in one region and little in another; the bovine-to-equine ratio also varied across regions). Parameters subject to larger amounts of variation are more likely to substantially influence model results more than those subject to lesser amounts of variation. When determining the

overall probability of introduction, we thus chose to focus our sensitivity analysis on the combinations of region and time period that were associated with the highest levels of risk and/or uncertainty.

3. Results

3.1. Spatiotemporal probability of entry

The overall annual median probability of EE entry in France was much higher than that for AHS for both pathways and for each of the three years considered (EE: PW-host = 0.9 and PW-vector = 0.4 to 0.5, AHS: PW-host = 3×10^{-3} and PW-vector = 1.4×10^{-2} to 3.6×10^{-2}). The route most susceptible to virus entry differed between diseases: for AHS, infectious vectors represented the route by which entry into France was most likely, while for EE, infectious hosts represented the most at-risk route of entry (see figure 3).

Seasonal effects were similar for both diseases, with a lower-risk period from January to June. This was the result from the assumption fact that the low- and very-low-risk regions (i.e. the exporting regions) were considered to be unlikely to experience AHS and EE outbreaks during this time and that animal import from high-risk regions is very rare in general (see Figure 3).

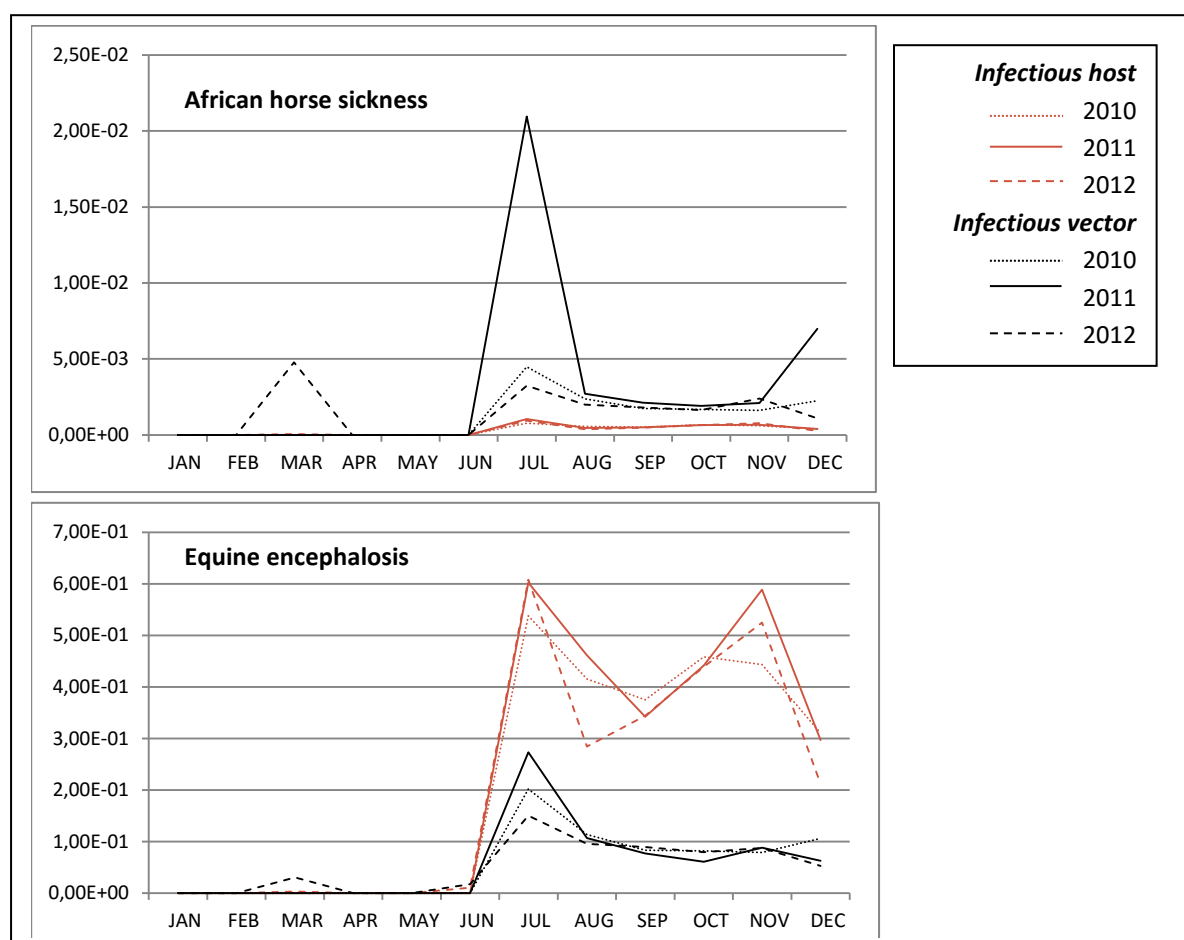


Figure 3 : National median probability of EE and AHS viruses entry to France.

Probability of entry varied greatly across space and time (see for example Figure 4 and Figure 5) but, over the three years considered, some areas consistently had a higher probability of virus entry, e.g., the southern and north-western regions of France. These areas were the most at-risk for both diseases and both entry pathways (see Figure 6).

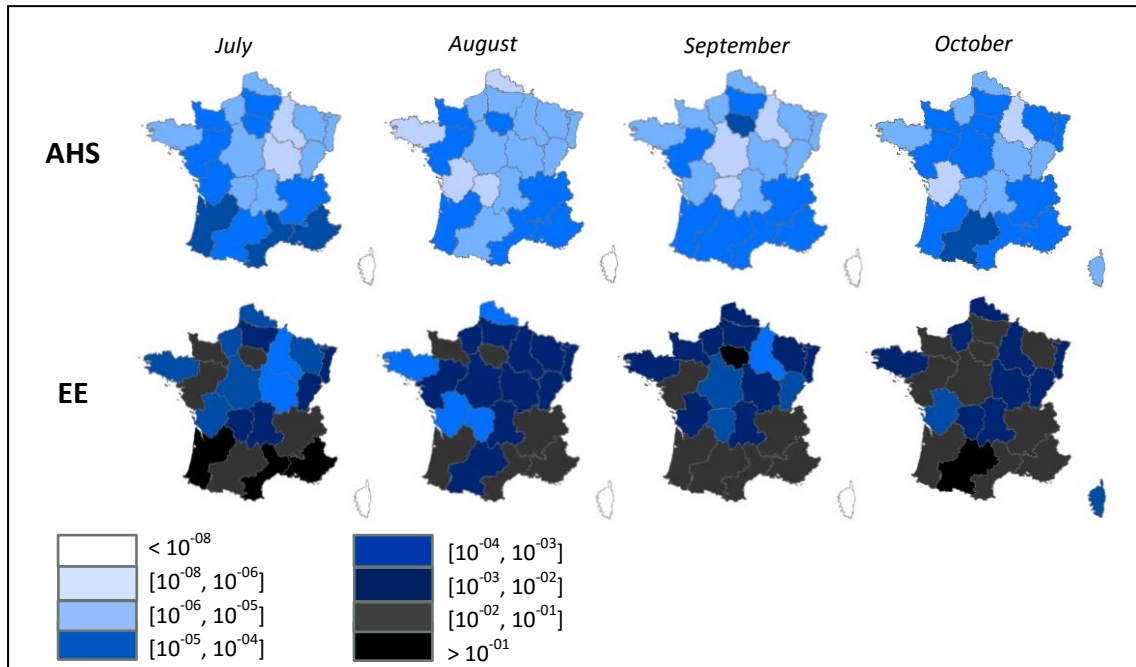


Figure 4 : Median probabilities of viruses entry via an infectious host. Example of the year 2012.

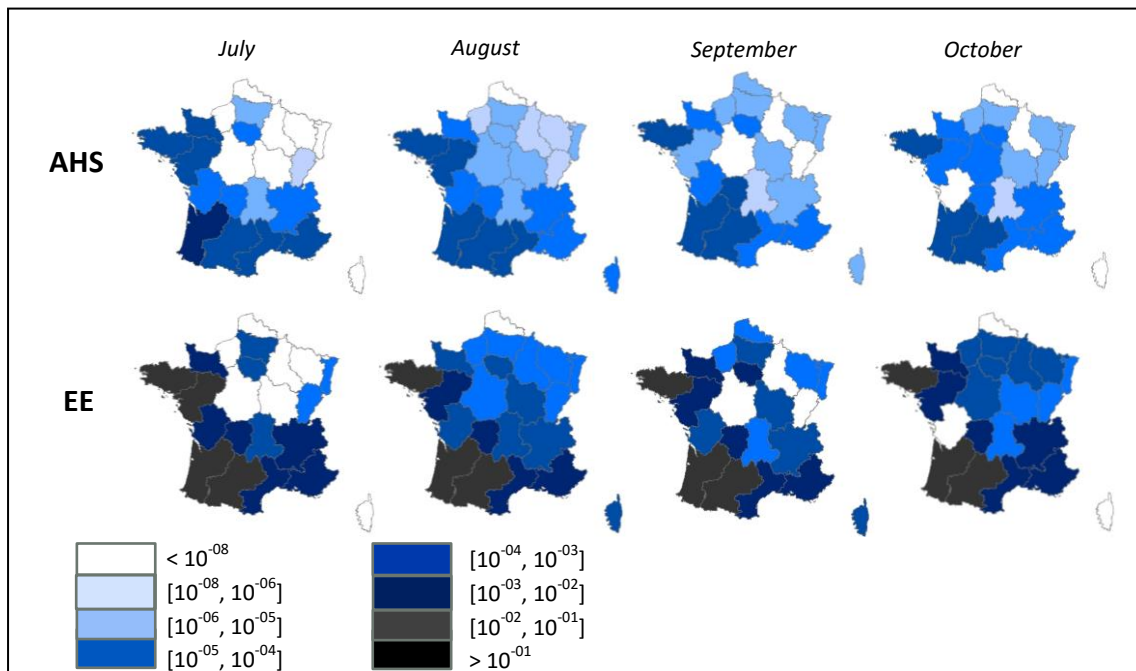


Figure 5 : Median probabilities of viruses entry via an infectious vector. Example of the year 2012.

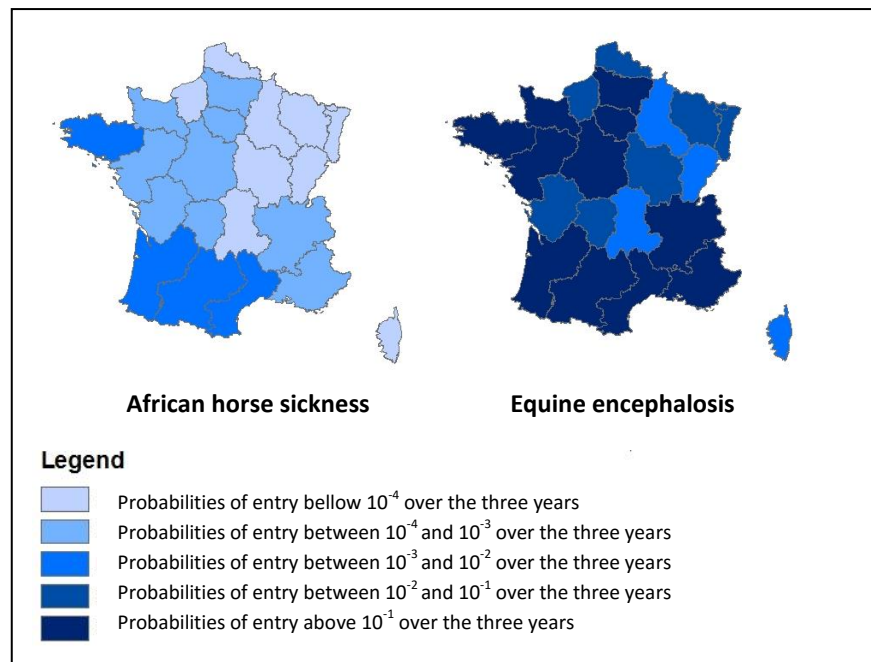


Figure 6 : At risk areas for viruses entry when entry via an infectious host and via an infectious vector are considered together.

3.2. Sensitivity analysis

The parameters with the greatest impact on the output differed between diseases and pathways. A summary of the input parameters with the greatest influence on the results is presented in Figure 7 and Figure 8.

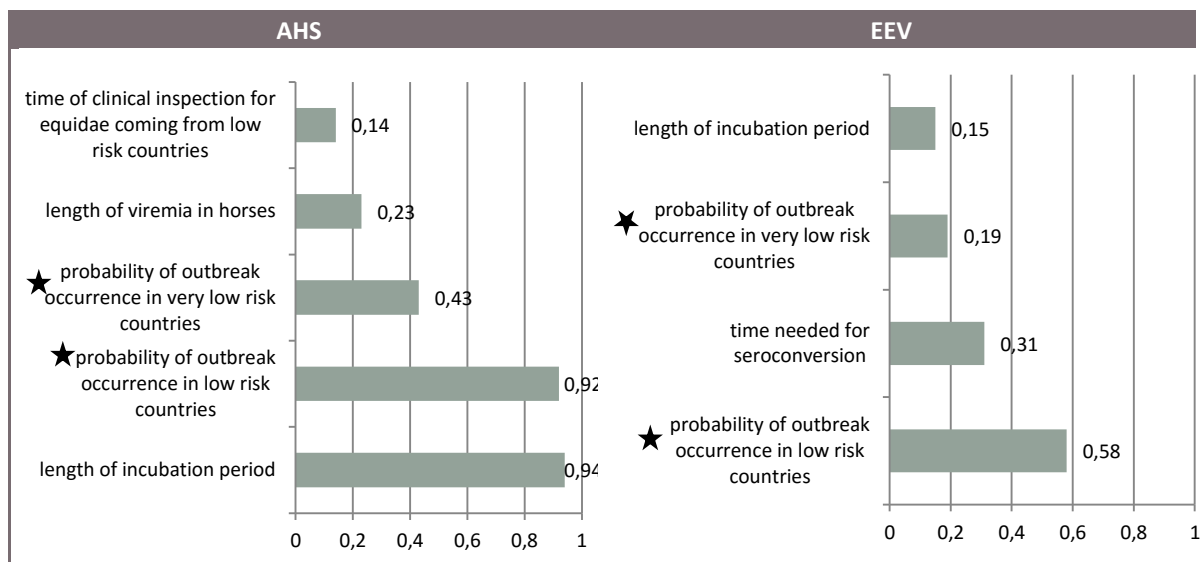


Figure 7 : Correlation of the model input parameters with the probability of entry of AHS and EE via an infectious host. Results are presented for Basse Normandie in July 2012. Only input parameters with at least one correlation $\geq |0.1|$ have been included in the table. Parameters with black star are the uncertain parameters. The others are the variable parameters.

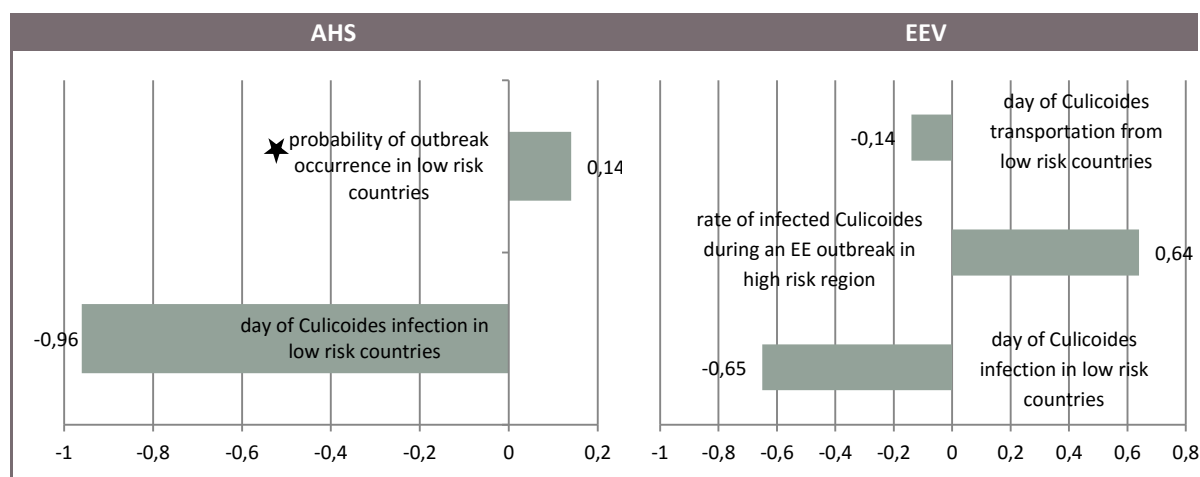


Figure 8 : Correlation of the model input parameters with the probability of entry of AHS and EE via an infectious vector. Results are presented for Basse Normandie in July 2012. Only input parameters with at least one correlation $\geq |0.1|$ have been included in the table. Parameters with black star are the uncertain parameters. The others are the variable parameters.

3.3. What-if scenarios

The what-if scenarios tested had similar effects on both diseases when the routes of introduction were considered separately (see Table 6). For example, the probability of importing an infectious vector decreased when systematic vector control was implemented during quarantine (a change of more than -40% for both diseases). Instead, changes to the length of the quarantine period and the use, or not, of laboratory tests had only a minor impact on the probability of introducing an infectious host (and obviously no impact on the importation of an infectious vector). Although the impact of the what-if scenarios was broadly similar (again, when considering each route of introduction separately) one difference should however be pointed out: a strategy of prohibiting imports from high-risk regions was much more successful in controlling AHS risk than EE risk. Indeed, in this scenario, the probability of importing an infectious vector decreased by 31.67% for AHS and by 3.40% for EE, and the probability of importing an infectious host decreased by 3.53% for AHS and 0.03% for EE.

When both entry pathways were considered together, the impact of the what-if scenarios on the overall probability of disease entry differed between AHS and EE. For example, when testing regulations were implemented in all low-risk countries, the overall probability of AHS introduction decreased by 4.99% while the probability of EE introduction decreased by 15.35%. Similarly, the systematic implementation of vector control on animals coming from low-risk countries had a greater impact on AHS probability of entry than on EE probability of entry (-46.13% for AHS and only -2.07% for EE). The most influential protective measures were thus disease-specific. To reduce the probability of AHS introduction, the most efficient protective measure was to implement vector protection on all animals coming from low-risk regions. To reduce the probability of EE introduction, the most effective measure was to implement the same regulation in all low-risk regions: quarantine before import for horses coming from both EU and non-EU countries, which resulted in a 15.36% reduction in the probability of disease entry.

Table 6 : Change (in %) on median probabilities of viruses entry to France compared to the default scenario.

	African horse sickness			Equine encephalosis		
	Infectious host	Infectious vector	Overall risk	Infectious host	Infectious vector	Overall risk
Quarantine period of 60 days instead of 40	-1.74	0	-0.17	-0.3	0	-0.10
Quarantine period of 20 days instead of 40	+2.69	0	+0.27	+0.08	0	+0.03
Same regulation implemented in Low risk EU state than in low risk non-EU state	-49.31	0	-4.90	-46.26	0	-15.36
No serological test implemented	+0.20	0	+0.02	+0.03	0	0.00
Vector protection implemented on all animals coming from low risk area	-2.87	- 51.07	-46.13	-0.35	- 57.26	-2.07
No importation from high risk region	-3.53	-31.67	-28.78	-0.03	-3.40	-0.12
No importation from low risk region and non-EU states	-2.88	-0.45	-0.69	-0.29	-0.70	-0.12

4. Discussion

Our study showed that the probabilities of entry for AHS and EE into France are similar in terms of seasonality and, in both cases, the highest-risk period is from July to December. This is mainly due to a decrease in the estimated risk from exports from low-risk countries from January to June, when the probability of having an outbreak was calculated as negligible (outbreaks never reported). The regions most susceptible to AHS and EE entry were also similar, with the highest probabilities of virus entry in the southern and northwestern regions of France.

However, the probabilities of AHS and EE entry differed in magnitude, with the latter much higher than the former. In addition, the main probability contributors also differed between diseases. For AHS, the most important pathway for virus entry was through an infectious vector, but for EE, the appearance of an infectious host was the most important contributor to the overall probability. These patterns were due to the less-strict trade regulations concerning EE (i.e. no serological tests in low-risk non-EU states), the higher probability of EE occurrence in low-risk regions (mean value 0.14) compared to AHS (mean value 0.016), the difficulties of detecting EE symptoms during clinical inspection, and longer high-risk periods for EE, which can be explained by the longer asymptomatic period of EE, during which time the disease is not reported in the exporting country. These aspects together explain why preventive measures for high risk areas are more effective for AHS than for EE.

The sensitivity analysis revealed that the fraction of infectious *Culicoides* during an EE outbreak in high-risk regions had a significant impact on EE entry via increasing the number of infectious vectors. This parameter was estimated as a constant in the AHS model due to the limited information available and thus did not appear in the sensitivity analysis. However, this result for EE suggests that the impact of this parameter would be not negligible for AHS entry. Considering the probability of virus entry via an infectious host, the time for seroconversion was highlighted as an important input parameter for EE (coeff. correlation equal to 0.31), but not for AHS (coeff. correlation equal to 0.01). This result can be explained by the increased relative importance of serological tests in detecting EE, as this disease is mainly asymptomatic. In addition, there was much more uncertainty regarding estimates of the time for seroconversion for EE than there was for AHS. This also contributed to the importance of this parameter in the EE model.

The what-if scenarios tested here showed that the most effective protective measures were not the same for the two diseases, because the main probability contributors differed. Thus, even if the impact of each scenario was similar between diseases for a given route of introduction, its impact on the overall probability of disease introduction varied. To specifically decrease the probability of importing an infectious host, the most efficient measure for both diseases was to implement the same regulations in all low-risk countries, EU and non-EU states alike. This pattern can be explained by the fact that relatively few horses are imported into France from either high-risk or low-risk non-EU states (respectively 0.08% and 2.8% of the total number of imported horses to France) compared to the number of horses imported from low-risk EU states (34.4% of total imports to France). These results were consistent with those obtained by de Vos and colleagues (de Vos et al. 2012). To specifically decrease the probability of importing an infectious vector, the most efficient measure was, as expected, the implementation of systematic vector control during quarantine, which is currently only mandatory for horses coming from high-risk areas.

The model allowed us to distinguish which disease posed the greatest threat in a given time and place even though, *a priori*, the diseases are similar in term of biological origin and epidemiological patterns. Our method also clearly identified the main probability contributors and the most efficient measures for risk mitigation. However, it does not take into account the assessment of the consequences of each disease, which would be useful in objectively allocating the limited resources for disease surveillance. For example, the probability of EE entry appears higher than that of AHS, but the consequences of AHS introduction would be more severe, given the pathogenicity of the disease. To aid in the allocation of resources, then, other approaches should be considered (e.g., (Cardoen et al. 2009; Havelaar et al. 2010; Krause 2008; McKenzie et al. 2007; Ruzante et al. 2010; Valenciano 2001)).

D. DISCUSSION AND CONCLUSION

1. Discussion

The approach used in this chapter provided a complete and detailed picture of the probability of having an outbreak. In particular, the spatiotemporal risk analysis highlighted important spatiotemporal variations in probability of introduction. This is consistent with the non-homogeneity of the horse population and movements and with the link between VBDs and climate/environment. Performing a spatiotemporal analysis is thus useful for identifying the regions and time periods that are most at risk. In addition, by combining the probability of entry and probability of establishment for two routes of virus introduction, we provided a more complete picture of the risk posed by a pathogen compared to risk assessments that involve only a single route of introduction or a single probability. The risk maps provided in section B (figure 3 in the paper) present a good illustration of these differences: the spatiotemporal probabilities are different according to the probability of release/entry, the probability of establishment, the probability of introduction, and the routes of virus introduction under consideration.

Our method can be easily adapted to other *Culicoides*-borne diseases, as illustrated by the work conducted on equine encephalosis (section C). However, the method can also be easily adapted to other vector-borne diseases as long as vector-related data (extrinsic incubation period, survival rate, abundance, etc.) and host-related data (viremia, incubation period, etc.) are available (keeping in mind, of course, that more adaptations would be required for models of diseases that are less similar to AHS). In particular, the probability of virus establishment in our example was quite simple to estimate, but could be more complicated for diseases with more complex transmission cycles (e.g., diseases with vertical transmission, multiple hosts and vector species, and/or vectors with a slow biological cycle, such as ticks). Careful consideration must thus be taken in adapting the model to another disease. This is especially true given our finding that even very similar diseases, such as EE and AHS, lead to quite different risk assessment outputs (see section C).

Despite its advantages, the approach applied in this chapter is still rarely implemented in animal health. The low number of publications found in the literature that combine routes of pathogen entry and probabilities of entry and establishment can be explained by two factors. Firstly, the propagation of bias and uncertainty that results from combining multiple data sources (Hoffman and Hammonds 1994) can complicate the interpretation of results. Secondly, each process requires a different estimate of probability and different measures of risk mitigation, which some researchers prefer to keep separate (Stevens et al. 2009). Indeed, a large amount of data is needed to build such a complex spatiotemporal risk model. These data were hard to obtain, and not initially ready to use. When data are routinely collected, most of the time they are not formatted for use in the context of risk assessment, and thus an important step of data preprocessing is often needed. As an example, the TRACES database contains data on horse movements, but it was not straightforward to extract the relevant information. The method

implemented in this chapter is thus time- and data-consuming, which constitutes one of the major obstacles to its wider implementation in risk estimation. This limitation may also restrict efforts to update the model and its outputs.

2. Conclusion

The approach used in this chapter provided a complete and detailed picture of the probability of experiencing an outbreak. However, a low probability does not mean that an outbreak is not occurring, and risk assessments do not predict with certainty the origin of an infection. This approach merely gives an estimate of the likelihood and the most likely sources of an outbreak. To ensure the early detection of a newly introduced disease, it would also be useful to consider other approaches for estimating the probability of experiencing an outbreak.

CHAPTER III: SYNDROMIC SURVEILLANCE

This chapter explores the possibility of implementing a syndromic surveillance approach to assess the probability of occurrence of a newly introduced VBD in horses. There is no single commonly accepted methodology for the execution of syndromic surveillance. We thus started by summarizing current methods and definitions used in syndromic surveillance and, in particular, in veterinary syndromic surveillance (Chap III.A). Based on this short review, we developed a syndromic surveillance system for the early detection of West Nile virus (WNV) (Chap III.B). WNV was chosen because it is currently a major concern in Europe for both human and equine populations. The syndromic surveillance system was developed using nervous symptoms in horses that are known to be early indicators of WNV. The output of this syndromic surveillance system was simple to understand, but also potentially complicated to use when data were close to the alarm threshold. In addition, it was difficult to combine the output with other epidemiological knowledge such as disease seasonality or environmental risk factors, parameters which are fundamental when working with VBDs like WNV. Indeed, how can we interpret a small outbreak occurring within the vector season versus a large outbreak occurring outside the vector season? To address this question, we tested an application of Bayes' rules to syndromic surveillance (Chap III.C). The objective was firstly to use syndromic surveillance to provide a quantitative assessment of the probability that an outbreak is in progress, and secondly to be able to combine syndromic surveillance with other epidemiological knowledge. The advantages and drawbacks of both approaches (classical and Bayesian approaches) in the assessment of the probability of a current VBD outbreak are discussed in the final section (Chapter III.D).

A. OVERVIEW

This section aims to present current methods and definitions used in syndromic surveillance and, in particular, in veterinary syndromic surveillance.

1. Overall principle of syndromic surveillance

1.1. History

1.1.1. *Syndromic surveillance and human health*

Syndromic approaches first gained momentum in human health applications, when the bioterrorist anthrax attacks of 2001 in the USA drew attention to the need for early detection of pathogen introduction (Buehler et al. 2003; Nordin et al. 2005). Concomitant outbreaks of new emerging infectious diseases, such as West Nile virus in 1999 (Henderson et al. 2001) and SARS in 2002 (Abdullah et al. 2003), reinforced the necessity of developing more timely surveillance systems. Real-time surveillance systems were then developed based on the automatic collection and transmission of pre-diagnostic and unspecific data, under the primary assumption that the behavior of these data change when a population's health is affected (Mandl et al. 2004a).

Syndromic surveillance was thus first used to enhance traditional passive surveillance, which is ineffective in detecting rare or emerging diseases due to the limited ability of clinicians to recognize the signs of unknown, or poorly known, diseases (Shaffer 2007). However, a syndromic approach is now also applied in the monitoring of well-known diseases such as human flu (Ginsberg et al. 2009; Hiller et al. 2013) in order to implement protective measures early and limit the impact of the disease.

1.1.2. *Syndromic surveillance and animal health*

In veterinary medicine, the development of syndromic surveillance systems followed a parallel path to that in human medicine. In particular, the recent focus on the 'One medicine' concept has contributed to increased awareness of early disease detection in animal populations (Dórea et al. 2011). However, compared to human medicine, syndromic surveillance in veterinary medicine poses its own unique problems. The great diversity of animal species and the types of production that must be considered constitute an impediment to the development of syndromic surveillance systems due to the lack of common vocabulary, practices, and data collection systems (Shephard 2006). In addition, animal data are subject to more variation stemming from non-disease factors, as a decision to seek care for livestock is mainly driven by a cost/benefit relationship and not, as in human medicine, by disease severity (Kosmider et al. 2006). Moreover, data regarding animal health are still scarce due to less frequent data collection and less developed data standards. All of these elements complicate efforts to monitor animal diseases.

Although the number of syndromic surveillance systems or initiatives is increasing (at least 27 systems were identified in 12 European countries in 2013 (Dupuy et al. 2013a)), fully operational and validated systems are still rare in veterinary medicine (Dórea et al. 2011).

1.2. Definition and objectives

The term ‘syndromic surveillance’ derives from the word ‘syndrome’, a set of clinical signs and symptoms that are correlated with each other. However, any data can be used, provided they are sensitive to changes in disease incidence in a population and contain an early signature of a disease outbreak.

Syndromic surveillance can thus use either real syndromic data (e.g., data from emergency departments (Hiller et al. 2013), estimates of mortality in cattle (Perrin et al. 2010)) or other health-related data (e.g., milk yield (Madouasse et al. 2013), search query data on Google (Dugas et al. 2013; Dugas et al. 2012; Ginsberg et al. 2009) and Twitter (Gesualdo et al. 2013; Signorini et al. 2011)).

The Triple-S project defines syndromic surveillance as “the real-time (or near real-time) collection, analysis, interpretation and dissemination of health related data to enable the early identification of the impact (or absence of impact) of potential human or veterinary public health threats which require effective public health action” (Triple S Project 2011). Beyond this, there is no single and commonly accepted definition for syndromic surveillance. The main underlying objectives can nonetheless be summarized with the following points (Katz et al. 2011):

- Early detection of and response to an outbreak, or at least detection of a probability of an outbreak high enough to warrant further investigation;
- Use of continuously acquired pre-diagnostic information;
- Possible applications during an outbreak, through the provision of tools for following the course of outbreaks;
- Providing assurance that an outbreak is not in progress;

The primary objective of all syndromic surveillance systems is thus to detect the signal, even a weak one, of an outbreak prior to its formal diagnosis.

1.3. Overall approach

There is no commonly accepted method or framework for the implementation of a syndromic surveillance system, and different approaches can be used. However, following preliminary steps, three main steps should be always implemented: 1) description and preprocessing of the data, 2) choice and implementation of an appropriate detection algorithm, and 3) evaluation of the system’s performance. Figure 9 summarizes the overall process of implementing a syndromic surveillance system. Each component is detailed in subsequent sections.

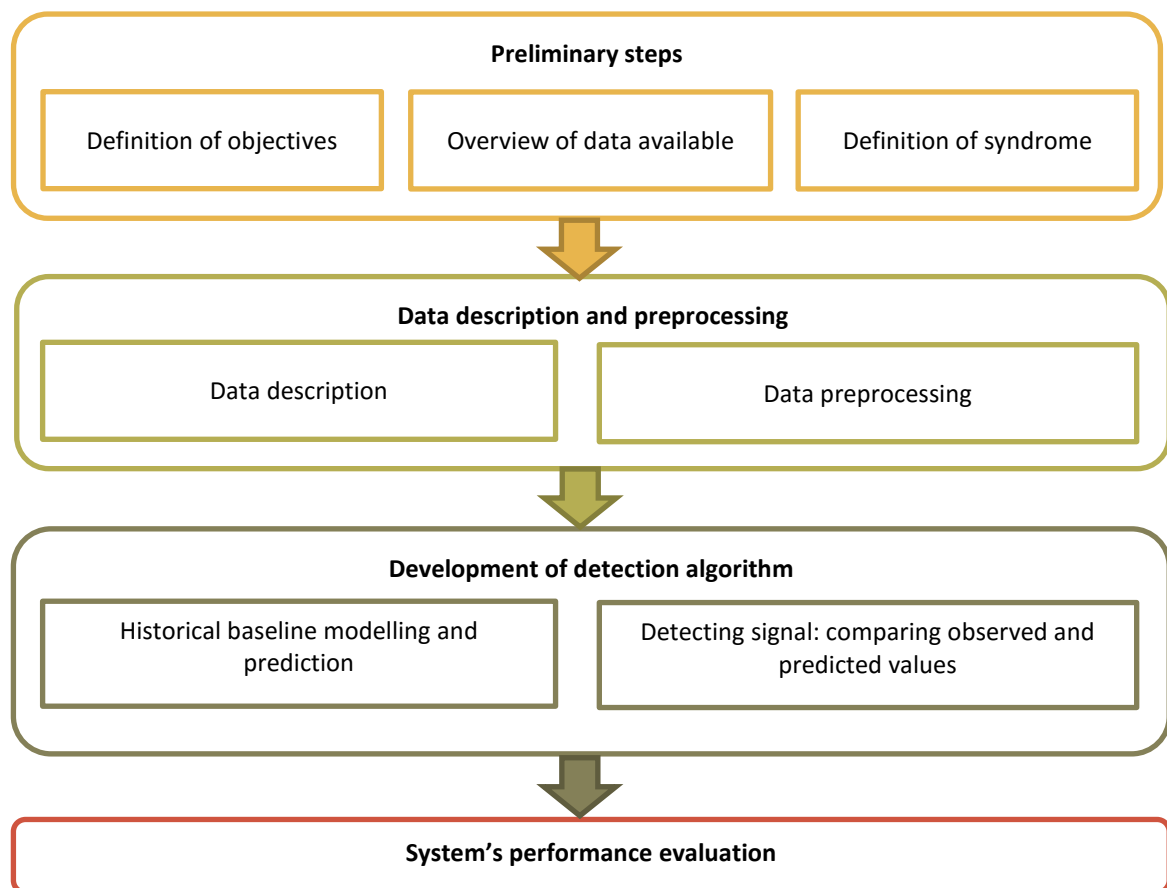


Figure 9 : Overall method to implement syndromic surveillance.

2. Preliminary steps

2.1. Definition of objectives

Like for all surveillance system, the first step of any syndromic surveillance system is to set clear objectives (i.e., disease(s) of interest, system users, desired balance between the sensitivity, specificity, and timeliness of detection). The geographic, demographic, and temporal coverage of the system must be also carefully assessed (Vial et Berezowski 2014). This step is essential as it will influence the performance of the surveillance system and its future organization.

2.2. Overview of data available

The essential data needed to achieve all surveillance objectives while minimizing the amount of data collected should be defined (Vial et Berezowski 2014). Then, as one of the key objective of syndromic surveillance system is to use of continuously acquired pre-diagnostic information, an inventory of the data sources available has to be made and evaluated in order to identify the data which can be used by the system.

2.3. Definition of syndrome

Then, it is important to set a clear definition for the group of clinical signs, or any other data considered, that constitutes an event of interest. This step is essential and has to be carefully considered, as it will influence the performance of the surveillance system (Ivanov et al. 2002; Shaffer 2007 ; Vial et Berezowski 2014). Various methods can be used for data classification. For example, Dupuy and colleagues (Dupuy et al. 2013b) used multiple factor analysis, while Dórea and colleagues (Dórea et al. 2013b) used naïve Bayes learning and decision trees.

This preliminary step poses unique problems in veterinary syndromic surveillance compared to human medicine. In particular, standards for data classification are not unified in veterinary medicine and each veterinary syndromic system develops and validates its own classification system (Dórea et al. 2011). This lack of standardized data constitutes a challenge in the definition of syndrome groups and of the rules for characterizing events of interest.

3. Data description and preprocessing

3.1. Data description

After the definitions of the syndrome and events have been determined, the data must be carefully analyzed in order to identify their main variations and characteristics. Potential aberrations due to past outbreaks or other events must be identified. The data must also be decomposed in order to identify systematic and stochastic variations. When considering time series, one must take into account systematic variations due to trends (long-term movements in the time series) or seasonality (cyclic variations). This preliminary work is important in order to determine which detection algorithm will be best-adapted to the data and if data preprocessing will be necessary (see Table 7).

3.2. Data preprocessing

Depending on the results of data decomposition, data can also be preprocessed. Two main examples of data preprocessing are presented: removal of aberrations and stationarity transformation.

3.2.1. Aberrations removal

As a result of past outbreaks or other events, raw data may contain unusual values (to which we refer as “aberrations”) which can disturb data modeling. To improve data modeling, and thus outbreak detection, these aberrations can be removed, either manually, based on historical data from previous outbreaks, or automatically, based on implemented procedures.

In particular, considering the specific case of time series data, Tsui and colleagues (Tsui et al. 2001) proposed removing data points above the 95% confidence interval of model predictions. This was based on the assumption that, after fitting the entire dataset to a regression model, data points above the 95% confidence interval of model predictions represent data from an epidemic.

Another option is to use a reweighting procedure to reduce the influence of high baseline counts, as proposed by Farrington and colleagues (Farrington et al. 1996). A weighting function is used on empirical grounds to assign very low weights to counts with large residuals. The residuals (s_i) are obtained based on the initial estimates (μ_i) and the dispersion parameters (ϕ). The weight of each value (w_i) equals a constant γ . If s_i is above 1, and $\gamma \cdot s_i^{-2}$ if s_i is below 1, the weighted data are refitted.

3.2.2. Stationarity transformation

The detection of outbreaks can sometimes be implemented only on stochastic variations of the data. This is especially true for time series which can be transformed into a stationary process. Such a transformation is essential in order to implement certain detection algorithms, such as control charts. In a stationary process, systematic variations have been removed from the data in order to keep only the stochastic variations.

4. Detection algorithms

There is no commonly accepted classification for algorithms used to detect abnormal patterns, but they are often presented according to the nature of the clusters they identify, i.e. temporal, spatial, or spatiotemporal clusters (Buckeridge et al. 2008). As our work centers on equine populations, in which spatial information is rarely available and/or accurate, we focus in this section on temporal methods for cluster detection. Following a summary of how to choose the most appropriate detection algorithm, we briefly present the algorithms commonly used in the detection of temporal clusters.

4.1. Choice of detection algorithm

The choice of the algorithm for detecting a signal is fundamental in determining the quality of detection. The choice must be based on (Dórea et al. 2011):

- the type of data available: single or multiple time series, data monitoring with rates or counts, long-term historical data available or not;
- the nature of the disease considered: sudden or slow increase in the number of cases;
- the desired balance between the sensitivity, specificity, and timeliness of detection.

The main detection algorithms used in temporal cluster detection are historical limits, control charts, and regression methods. A recent review showed that regression methods were the most popular methods

used in veterinary syndromic surveillance in Europe (Dupuy et al. 2013a). Indeed, among the 23 systems found, eight used regression methods (of which four used autoregressive moving average models (referred to as “time series methods”) and the remaining four used other methods), four systems used historical limits, and only one used control charts. Control charts are, however, more frequently used outside Europe (Dórea et al. 2011).

The popularity of regression methods might be explained by their ability to easily deal with data involving trends or seasons. Nevertheless, the drawback is that these methods require a long historical baseline, which is often not available. Conversely, control charts do not require a long historical baseline, but they also require preconditioning to remove seasonality, trends, and other variations, which adds complexity to the analysis. Historical limits algorithms are the simplest methods but they suffer from numerous drawbacks which may compromise outbreak detection (e.g., no adjustment for trends or disease clusters). A summary of the main advantages and limitations of these different detection algorithms is presented in Table 7.

Table 7 : Summary of the advantages and limitations of the main detection algorithms in the time series analysis.

Detection algorithm	Advantages	Limitations
Regression		
<i>ARMA</i>		Needs long historical baseline; requires preconditioning to remove seasonality, trends, and other variations
<i>ARIMA</i>	Trends taken into account	Needs long historical baseline
<i>SARIMA</i>	Seasons taken into account	Needs long historical baseline
<i>HW</i>	Needs short historical baseline; easily automatable; adaptable to local changes	Only one type of seasonality allowed
<i>lm and glm</i>	Explanatory variables present; several type of seasons possible	Needs long historical baseline
Historical limits	Easy to implement	Needs long historical baseline; no adjustment possible for trends, disease clusters, or aberrations; no consideration of reporting delays; inconsistent case inclusion criteria
Control chart		
<i>Shewhart chart</i>	Needs short historical baseline; peak-shaped outbreak	Requires preconditioning to remove seasonality, trends, and other variations; detection only based on the last previous value
<i>EWMA</i>	Needs short historical baseline; outbreak in the form of a slow increase	Requires preconditioning to remove seasonality, trends, and other variations
<i>CUSUM</i>	Needs short historical baseline; outbreak in the form of an increase longer than one time unit	Requires preconditioning to remove seasonality, trends, and other variations

4.2. Historical limits

One simple method to detect outbreaks is based on historical limits and was first applied in the United States by Stroup and colleagues (Stroup et al. 1989).

An alarm is triggered when:

$$\frac{X_0}{\mu} > 1 + 2 \times \frac{\sigma_x}{\mu}$$

where X_0 is the number of cases in the most recent four-week interval and μ and σ_x are the mean and the standard deviation, respectively, of the previous 15 historical four-week periods (from X_1 to X_{15}).

This method is simple to implement but presents major bias due to a lack of adjustment for gradual trends or disease clusters, a lack of consideration for reporting delays, and inconsistency in case inclusion criteria between current and historical data. Although a method for bias reduction was recently proposed (Levin-Rector et al. 2015), no real evaluation of the detection performance of a historical limits algorithm was conducted.

4.3. Control charts

4.3.1. Principle

Control charts are based on graphical representation of data and are commonly used when no solid baseline is available. Control charts rely on cumulative differences between observed data in a time window and a threshold (Mandl et al. 2004a), which is set at some multiple of the standard deviation of the sample's mean (Carpenter 2002; Hutwagner et al. 2005). The value of the multiplier is chosen to optimize tradeoffs between sensitivity, specificity, and timeliness in order to meet the objectives of surveillance. However, the assumptions behind these methods are that the data are independent and distributed according to a known hypothesis, assumptions that are not met by most surveillance data (Lotze et al. 2008). In order to use control charts, then, significant preconditioning must be applied to transform the data into a stationary process.

4.3.2. Methods

Different kinds of control charts are available: Shewhart charts, cumulative sums (CUSUM), and exponential weighted moving average (EWMA) methods.

Shewhart charts are appropriate when a single peak is expected, and are simply based on the difference between observed values and average values calculated within a moving window (Shewhart 1931). Only the last mean recorded value is used for anomaly detection.

The CUSUM and EWMA methods rely on cumulative differences between observed data in a time window and a threshold (Mandl et al. 2004a). They are better-adapted to detect small but repeated lags between observed and expected values. Model parameters are empirically set to the best balance between sensitivity, timeliness, and specificity that is chosen to meet the objectives of the surveillance.

The CUSUM method is based on the calculation of a cumulative asymmetric sum and is especially useful in detecting outbreaks that are longer than one time unit (O'Brien and Christie 1997). It is sensitive to small shifts in deviations from the mean and also detects deviations more quickly than some other statistical process control methods. Two examples of the use of CUSUM are presented in (Maciejewski et al. 2007; Shaffer 2007), who both used CUSUM methods to detect syndromic aberrations in pets. This method is based on the equation:

$$C_t = \max \{0, (D_t + C_{t-1})\}$$

where t is the current time point and D_t is the standardized difference between the current observed value and the expected value. The differences are accumulated daily (for each point t , the statistic incorporates the value at point $t-1$) over the baseline but reset to zero when the standardized value is negative.

The EWMA method is based on exponential smoothing and is especially useful in the detection of a gradual increase in a signal (Hunter 1986). One example of the application of EWMA methods to the detection of aberrations in veterinary laboratory data is presented in (Dórea et al. 2013a). EWMA is based on the equation:

$$E_t = (1 - \lambda)^t E_0 + \sum_{i=1}^t (1 - \lambda)^{t-i} \lambda I_i$$

where λ is the smoothing parameter (>0) that determines the relative weight of current data to past data, I_t is the observed value at time t , and E_0 is the starting value.

4.4. Regression methods

4.4.1. Principle

Regression methods can be used when long-term historical data are available (e.g., (Kosmider et al. 2006); (Dórea et al.); (Perrin et al. 2010)). They utilize statistical methods for fitting a model to observed data in order to make predictions. The benefit of such an approach is that seasonal effects and trends that are observed in the dataset can be readily incorporated into the equation. The models are usually assessed based on an analysis of residuals and goodness-of-fit, and compared using AIC and root-mean-squared error within and outside the calibration period. In veterinary syndromic surveillance, these methods are still mainly used to perform retrospective analyses in order to assess their potential for prospective modeling (Dórea et al. 2011).

4.4.2. Regression models

Typical models are regression-type models (e.g., generalized linear models), autoregressive moving average (ARMA) models, and exponential smoothing. The different regression models are briefly presented in this section.

- **Linear and Generalized linear models**

Linear models (LMs) and generalized linear models (GLMs) are common tools for fitting data. In syndromic surveillance, they are especially useful when time series show trends and/or seasonal variations and when explanatory variables are used. Indeed, these types of information can be directly included in the model without additional data preprocessing. Two examples of their use in veterinary syndromic surveillance are shown in (Perrin et al. 2010) and (Kosmider et al. 2006), who applied Poisson regression on cattle-related data (mortality and number of laboratory isolations of *Salmonella*, respectively).

- **Autoregressive moving average models**

Autoregressive moving average models (ARMA) are based on two processes: one for the auto-regression (AR) and another for the moving average (MA) (Box et al. 2008). ARMA models require data to be stationary, i.e. without trends and with a mean and variance that do not change over time. If the data are not stationary, a common solution is to use a difference variable to transform the time series into a stationary process. Another option is to use an autoregressive integrated moving average (ARIMA) or seasonal ARIMA (SARIMA) model. The first aims to incorporate non-stationarity in the mean and the second attempts to take into account seasonal variations (Box et al. 2008).

- **Exponential smoothing**

Exponential smoothing involves exponentially decreasing the weights of observations over time, such that oldest observations have the smallest weight (Gardner 1985). The forecast is continuously revised according to more recent observations. The EWMA, or exponential weighted moving average, approach is the simplest form of exponential smoothing and is used when data do not have trends and/or seasonality. Instead, if trends are present, double exponential smoothing is used.

Triple exponential smoothing, also called Holt-Winters exponential smoothing (HW), aims to take into account both trends and seasonality. The seasonality can be either multiplicative or additive, but there can be only one type of seasonal pattern. If more than one kind of 'seasonal' pattern is present (e.g., monthly and daily seasonality), another smoothing method must be chosen. HW incorporates three components: a level term, a trend term, and a seasonality term, respectively defined by the smoothing constants α , β , and γ . The main advantage of this technique is that it is easily automatable and adaptable to local changes in the data (Lotze et al. 2008). One example of the use of HW in veterinary syndromic surveillance is found in (Dórea et al. 2013a).

4.4.3. Threshold values

The threshold values that trigger an alarm are typically a multiple of the standard error of the prediction (Mandl et al. 2004a), based on Serfling's approach (Serfling 1963). This constant is determined by the best compromise between sensitivity and specificity for a given case, as illustrated, for example, by Muscatello *et al.* (Muscatello et al. 2008). The authors explored excess mortality due to influenza and discussed the optimal threshold to balance the false positive alarm rate and true positive alarm rate. In general, a value between 2- and 3.5-times the standard error is often chosen to ensure the false alarm rate is below 5% (Mandl et al. 2004a).

Farrington and colleagues (Farrington et al. 1996) also proposed the use of an exceedance score based on the upper limit of the 99% prediction interval. An alarm is triggered when:

$$X_i = \frac{(\gamma - \mu_i)}{(U - \mu_i)} > 1$$

U_i is the upper limit of the 99% prediction interval, μ_i is the initial estimated value, and γ is a constant.

5. Assessment of performance

Assessment of syndromic surveillance systems is essential for determining the validity of conclusions. The assessment can be either qualitative or quantitative. However, in veterinary syndromic surveillance, assessment of the system's performance is still rarely carried out (Dórea et al. 2011). Moreover, when assessment is performed, it mainly focuses on data quality rather than real system performances (Dupuy et al. 2013a).

5.1. Performance metrics

Quantitative evaluation of a syndromic surveillance system's performance can be implemented in a similar way to the evaluation of the performance of individual diagnostic tests. Sensitivity, specificity, predictive positive values, and predictive negative values are commonly used as evaluation metrics of outbreak detection algorithms (Buckeridge et al. 2005; Choi et al. 2010). Regarding sensitivity, two approaches can be considered: the outbreak-day approach and the outbreak-detection approach (Mandl et al. 2004b). The outbreak-day approach considers each day (or time period considered) within an epidemic period to be an independent case. A true positive alarm is thus produced for each outbreak-day detected. The outbreak-detection approach considers each outbreak to be a single entity and thus, a true positive alarm is produced when at least one outbreak-day is detected. Various other quantitative parameters can be used to compare detection algorithms, according to the objectives of the surveillance system (see Table 8).

In their “Framework for evaluating public health surveillance systems for early detection of outbreaks”, Bühler *et al.* (Bühler et al. 2004) suggested a list of criteria, such as usefulness, flexibility, and acceptability, to qualitatively describe and evaluate each system component. Regarding veterinary syndromic surveillance, other criteria should be added, including population coverage, automation of data capture and transfer, benefit to users, detection efficiency of programmed algorithms, and contribution to claims of disease freedom (Dórea et al. 2011).

Table 8 : Metrics used for evaluation of outbreak detection algorithms
(Buckeridge et al. 2005, Choi et al. 2010)

Parameters	Definition
Sensitivity	Probability of alarm given that an outbreak is occurring. Can be based on overall outbreak detection or on the outbreak day number (each day considered a separate and independent case)
Specificity	Probability of no alarm given that an outbreak is not occurring
Predictive positive value	Probability that an alarm signals a true outbreak
Predictive negative value	Probability that no alarm corresponds to a true absence of an outbreak
ROC curve	Plots sensitivity (or true positive rate) against 1-Specificity (or false positive rate) for a range of algorithm settings
Area under ROC curve	Summarizes the detection performance of an algorithm. Values larger than 0.5 indicate that the algorithm is better than a random detection scheme.
AMOC curve	Plots a summary measure of time-to-alarm (given an outbreak is occurring) against the false positive rate
FROC curve	Plots the fraction of outbreak detected against the false positive detection rate
ARL	Expected time until the first detected event - ARL0; the expected time-to-alarm when there is an ongoing outbreak at the initiation of surveillance - ARL1
PSD	Probability of an alarm before some critical point in the outbreak given that the outbreak is detected
Time lag	Average number of weeks between the first of a consecutive number of outbreaks and the first alarm raised by each method
Missing rate	Number of missed outbreaks/total number of outbreaks

5.2. Test data

Generally, models are built using a calibration period and their predictions are then tested using a validation period. The role played by validation data in algorithm research is a crucial one and determines the validity of conclusions, especially when attempting a quantitative assessment of the system’s performance (Buckeridge et al. 2005). Authentic baseline and outbreak data can be used but the scarcity of such data means that it is often not possible to generate a quantitative assessment of the detection algorithm. Simulations of background data and/or outbreaks are therefore often required.

To simulate the background of a time series, different methods can be implemented. One of the simplest methods is to set the predicted value for each time period as the mean of a Poisson distribution. This

distribution is then sampled randomly to determine the value for a week of a given year, as proposed by Dórea et al. (Dórea et al. 2013a). For outbreak data, most studies have used naturally occurring signals from one or more real outbreaks to evaluate outbreak detection performance (Buckeridge 2007). However, given the limited availability of high-quality data from known outbreak periods, such signals can also be simulated. The methods for doing so range from simple mathematical functions to more complex stochastic simulation. Outbreak simulation has to consider the duration of an outbreak, spacing between outbreaks, temporal progression, outbreak magnitude, and spatial features (Mandl et al. 2004b). The advantages and disadvantages of each set of tests are presented in the Table 9.

Table 9 : Advantages and disadvantages of types of test data (Buckeridge et al. 2005)

Type of set	Advantages	Disadvantages
Wholly authentic	Face validity; authentic background and outbreak signals	Sufficient resources required to define outbreaks; validity and reliability of outbreak indications may be poor and difficult to assess; limited number and variety of outbreaks
Wholly simulated	Exact specification of outbreak signal; large number of test data possible; can be simple to develop; enables sensitivity analyses	Complexity of simulating baseline and outbreak signal; validity may be poor and difficult to assess; can require many parameter values
Simulated outbreaks superimposed onto authentic data	Greater face validity than wholly simulated test sets; exact specification of outbreak signal; large number of test sets possible; enables sensitivity analyses	Complexity of simulating outbreak signal; validity may be poor and difficult to assess; can require many parameter values

6. Conclusion

There is no single commonly accepted method or framework for the implementation of a syndromic surveillance system, and different approaches can be used to detect temporal clusters of abnormal events. However, three main steps should always be included: 1) description and preprocessing of the data, 2) choice of an appropriate detection algorithm, and 3) assessment of the system's performance.

Considering the use of syndromic surveillance in veterinary medicine, some specific points can be highlighted. Firstly, it can be challenging to define syndromes and abnormal events due to the lack of standards of data classification. Secondly, the most commonly used detection algorithms are control charts which, despite their limitations, are useful when a long-term baseline is not available. Regression methods are *a priori* more robust and flexible but are often only used for retrospective analyses because of the absence of long-term historical data. Finally, assessment of system performance in veterinary syndromic surveillance remains rare and, when it is performed, it mainly focuses on data quality rather than system performance.

B. COMPARISON OF PRE-PROCESSING METHODS

In this section, one of the above-described classical methods is applied to the detection of West Nile virus in horses. This part of the work resulted in a poster presentation at the Conference of the Society for Veterinary Epidemiology and Preventive Medicine (SVEPM) in Ghent (2015). The poster is available in Appendix 9.

1. Introduction

The passive French surveillance system 'RESPE' (introduced in Chapter I.C.1.3) has collected data on nervous symptoms observed in French horses since 2006. Although the sentinel veterinarians involved in this system are present in most French regions (92 of 96), the nervous symptoms collected are mainly reported from areas with high horse densities (see Figure 10). Diagnostic tests for West Nile virus (WNV), equine herpes virus serotype 1 (EHV-1), and consensus equine herpes virus (EHV-sp) (Léon et al. 2008) are systematically implemented for each declaration.

Currently, the collected data are mainly used to produce alerts when cases with positive laboratory diagnoses are identified. The data are also used for basic syndromic surveillance: an alarm is triggered when four syndromes are reported in the same week, or three declarations reported in each of two consecutive weeks. This alarm threshold was set arbitrarily and alarms may result in the initiation of epidemiological investigations depending on the context of the declarations. However, the reliability of this threshold has never been assessed and the ability of the RESPE nervous syndrome database to serve as a routine syndromic surveillance system is currently unknown.

Nervous syndromes in horses are considered to be an early indicator of WNV outbreaks (Leblond et al. 2007). Using routinely collected RESPE data in an early detection surveillance system could lead to the timelier implementation of protective measures, thereby limiting the consequences of a WNV outbreak for both equine and human populations. Therefore, there is a need to assess the capacities in which the RESPE nervous database can be used to detect WNV outbreaks.

The RESPE nervous syndrome dataset presents several aberrations that can be mainly explained by EHV-1 and EHV-sp outbreaks. As discussed in Chapter III.A, different methods can be used to remove these unusual data (manual or automatic procedures). To date, though, little work has been done on the potential impact of the choice of pre-processing method on a surveillance system's ultimate performance.

In this section, we use several preprocessing methods and detection algorithms to model time series data from the RESPE nervous symptoms database. The objectives were (i) to evaluate the impact of these methods and algorithms on system performance, and (ii) to assess whether or not these data can be used as a routine syndromic surveillance system for the detection of WNV outbreaks.

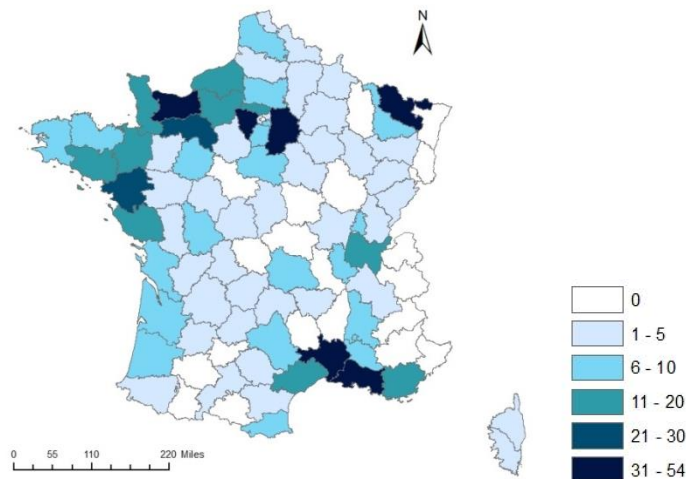


Figure 10 : Number of nervous symptoms cases declared to RESPE from 2006 to 2013.

2. Methods

2.1. Data characterization

In the RESPE database, nervous symptoms in horses are defined as any signs of impairment of the central nervous system, i.e. ataxia, paresis, paralysis and/or recumbency, and/or behavioral disorder. Cases, or an unusual cluster of cases, with “atypical” expression (colic, lameness, excitement, falling, muscular atrophy) can also be considered after the most common etiology of these symptoms has been excluded, since these signs can sometimes be the clinical manifestation of an affected central nervous system. Nervous disorders with evidence of traumatic or congenital origins are excluded.

Data on nervous symptoms in horses were available from RESPE for every calendar day from January 1st, 2006 to December 31st, 2013, totaling 532 declarations. An initial data characterization was performed using the daily data in order to identify explainable patterns such as global linear trends and seasonality. However, in the remainder of the study, the time series was aggregated into weekly counts due to the low per-day count. Monthly aggregation was not considered, as the main objective of this surveillance system was early detection.

Tests for WNV and EHV are routinely carried out on horses that present nervous symptoms, and the database contains 80 positive laboratory results, mainly for EHV-1 (only two positive cases of ELISA-IgG West Nile virus). The EHV-1 positive cases were either isolated cases – i.e. not associated with other positive cases – or from a cluster of cases that could represent a true outbreak.

2.2. Data pre-processing

We split the data into two time periods: data from 2006 to 2010 were used to train the model (see next section) and data from 2011 to 2013 were used to validate the model.

The raw time series used for model training was called TS0. We investigated three options for the removal of aberrations present in TS0 in order to obtain an outbreak-free baseline. In the first method, we retained only the 452 cases with no positive laboratory results (TS1). The second method consisted of removing all data linked to historical EHV-1 outbreaks, based on information from the RESPE website (TS2). This method did not remove single positive cases but only the positive cases associated with a cluster of other positive cases. In our third method, extreme values from TS0 were removed using the approach of Tsui and colleagues (2001)(Tsui et al. 2001), which assumes that, after the data have been fitted to a regression model, data points above the 95% confidence interval of the model prediction represent an outbreak (TS3). The authors used Serfling's regression model (Serfling 1963), which is a linear regression model that uses sine and cosine terms to account for seasonal variation. With our own data, we followed the proposal of Dórea and colleagues (Dórea et al.) and used a Poisson regression, which they considered an appropriate method to capture baseline activity while minimizing the influence of aberrations present in the dataset. The data were thus first fitted to a Poisson distribution and then values above the 95% confidence interval were removed. In TS1, TS2, and TS3, the values of weeks considered to be part of an outbreak were not removed but instead replaced by the average of the four previous weeks.

The four time series are shown in Figure 11.

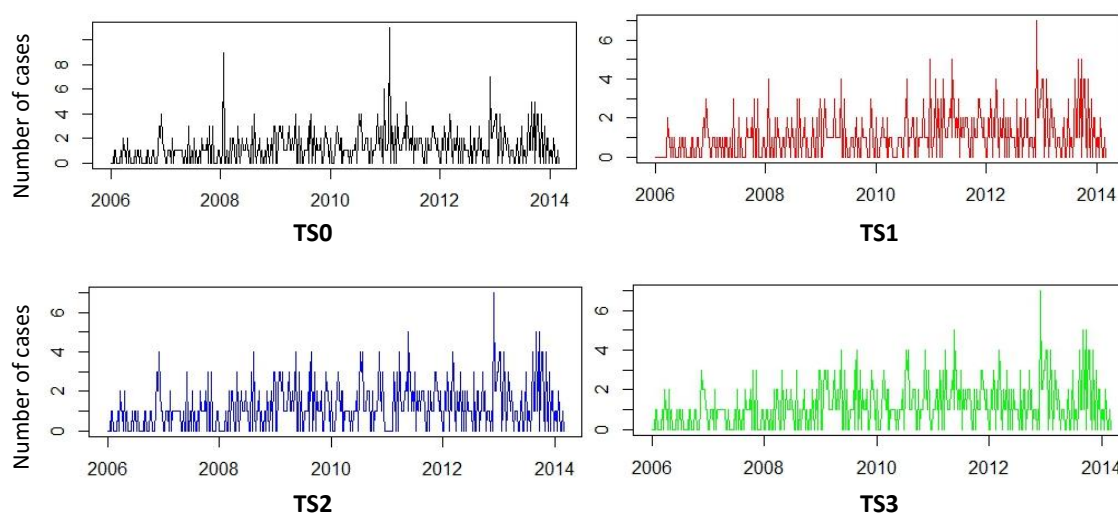


Figure 11 : Four time series used with datapreprocessing used from 2006 to 2010 and raw data used from 2011 to 2014. TS0 = raw data, TS1 = only the cases with no positive laboratory results, TS2 = outbreaks removed based on historical data, TS3 = extreme values above the 95% confidence interval deleted.

The explainable patterns (such as global linear trends and seasonality) were investigated in each time series (TS0, TS1, TS2, and TS3) in order to assess the impact of preprocessing methods on the dataset. We generated summary statistics by month and year, and performed moving average and autocorrelogram analysis (Lotze et al. 2008).

2.3. Forecasting

Forecasting was attempted using generalized linear regression models (GLMs) that were appropriate for count data (Poisson and negative binomial (NB) regressions) and Holt-Winters generalized exponential smoothing (HW). For GLMs the evaluated models included different types of seasonality through the use of sinod models with 1, 2, or 3 periods/year and season or month as factorial variables. To account for differences between years, we calculated the average counts for 53 consecutive weeks (*histmean*). To ensure that an ongoing outbreak would not influence the estimate, we used a 10-week guard band for the calculation of *histmean*.

Training data from 2006 to 2010 (TS0, TS1, TS2, TS3) were used to train the models, while data from 2011 to 2013 were used to validate the quality of the predictions. Alternative GLMs were evaluated for training data using the Akaike information criterion (AIC) (Bozdogan 1987). For HW, the optimal parameters were determined through minimization of the squared prediction error (Kalekar 2004).

The best models were then evaluated and compared using the autocorrelation and partial autocorrelation functions of the residuals (ACF and PACF, respectively) and the root-mean-squared error (RMSE). ACF and PACF are used to find repeating patterns (e.g., seasons) in a dataset. ACF is the linear dependence of a variable on itself at two points in time and PACF is the autocorrelation between two points in time after removing any linear dependence between them (Box et al. 2008). RMSE is a measure of the difference between the values predicted by a model and the values actually observed from the environment that is being modeled (Chai and Draxler 2014). This criterion was calculated for the differences between the observations and the predicted values within both the calibration period ($RMSE_c$) and the validation period ($RMSE_v$). In either case, the lower the criterion, the better the predictive performance of the model.

2.4. Detection algorithm

Finally, all eight combinations of pre-processing (4) and forecasting methods (2) were evaluated on their ability to detect simulated disease outbreaks: GLM applied to TS0, TS1, TS2, and TS3; and HW applied to TS0, TS1, TS2, and TS3. A six-week guard band was used to ensure that previous outbreaks would not influence the estimate of the baseline. The outbreak detection method used was based on a multiple of the standard error of the prediction. The action threshold was defined as the predicted number of cases in a given week plus a constant multiple of the standard error of the model prediction. If the actual observed value was above the threshold, an alarm was triggered. The constant multiple was empirically defined according to the best balance between sensitivity and specificity.

Baseline data from 2011, 2012, and 2013 were used for the assessment of the detection algorithms' performance. We simulated WNV outbreaks based on historical data from three previous European outbreaks: French outbreaks in 2000 (Murgue et al. 2001) and 2004 (Leblond et al. 2007) and an Italian

outbreak in 1998 (Autorino et al. 2002). The average weekly count of nervous-symptom cases in horses was calculated from the three historical outbreaks for an epidemic period covering a total of 11 weeks, from the first positive case detected to the last positive case detected (see Figure 12). The number of cases for each week of an epidemic period was sampled randomly between the extreme values obtained from historical data.

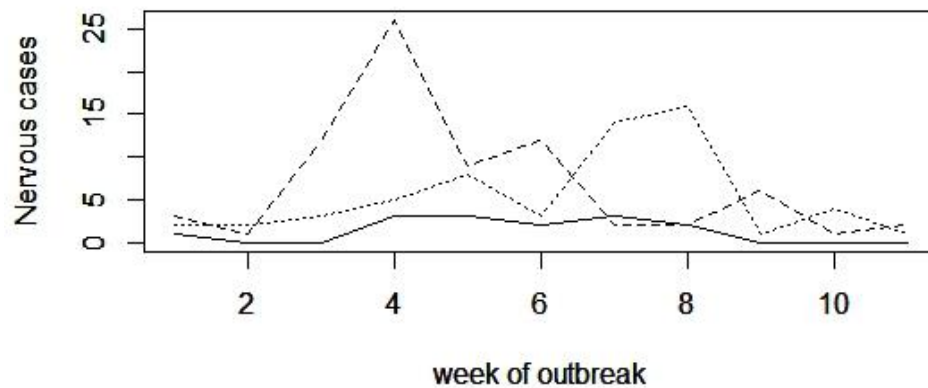


Figure 12 : West Nile virus outbreaks. solid line = outbreak Italy 1998 (Autorino et al. 2002), dotted line = outbreak in France 2004 (Leblond et al. 2007), dashed line = outbreak in France 2000 (Murgue et al. 2001).

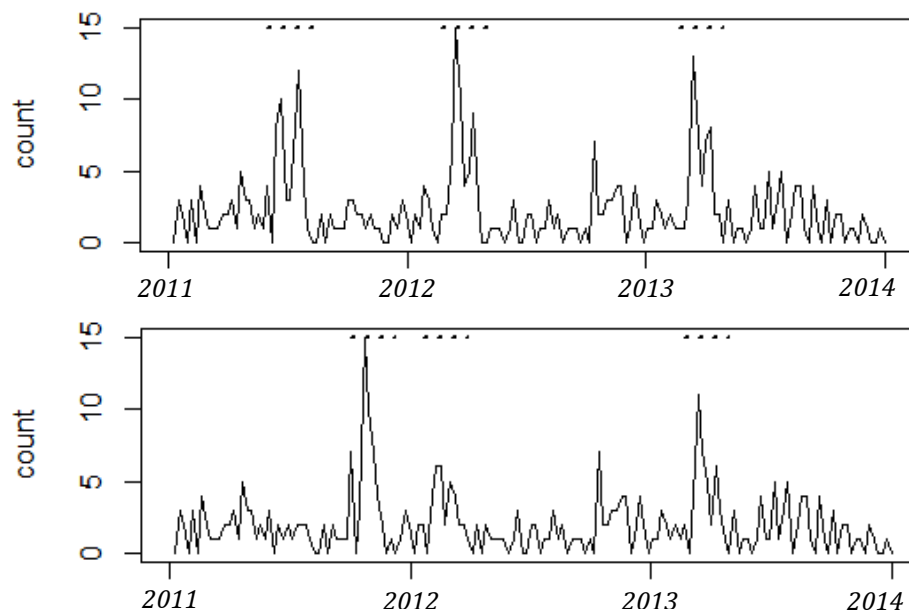


Figure 13 : Two examples of simulated outbreaks inserted in TS0. Positions of outbreaks identified with dotted lines above the peak.

To test our detection method, three simulated outbreaks were randomly inserted in the baseline data from 2011, 2012, and 2013, with at least 15 weeks between each outbreak in order to avoid overlap (see example in Figure 13). The process was repeated 25 times, for a total of 75 years containing a total of 75 outbreaks.

2.5. Quantitative assessment

We first calculated sensitivity based on the number of outbreaks detected out of all inserted outbreaks and denoted this Se_{out} . An outbreak was detected when it triggered at least one true alarm, defined as a week that produced an alarm and that was a part of an epidemic period. Se_{out} was calculated as:

$$Se_{out} = Out / (Out + No_Out)$$

where Out is the number of outbreaks detected and No_Out is the number of outbreaks not detected.

We also calculated Se_{wk} , the sensitivity based on the number of weeks in an epidemic period in which an alarm was triggered. Se_{wk} and specificity (Sp) were calculated as:

$$Se_{wk} = TP / (TP + FN)$$

$$Sp = TN / (TN + FP)$$

where TP is the number of true positive alarms, TN the number of true negative alarms, FP the number of false positive alarms, and FN the number of false negative alarms.

A receiver operating characteristic (ROC) curve was generated in R by testing various alarm thresholds, and the area under each curve (AUC) was also calculated (Hanley and McNeil 1982). The time to the first true alarm within an epidemic period was also evaluated in order to assess the efficiency of early detection.

2.6. Implementation

Models were implemented in R x64 version 3.0.2. Dynamic regression was performed with the functions *glm* (package {stats}), *glm.nb* (package {MASS}), and *stlf* (package {forecast}). The expected numbers of counts at time t were estimated with the *predict* functions of the respective packages. The expected numbers of outbreak-related cases were estimated with the *fitdist* function of the package {fitdistrplus}. AUCs were estimated with the *auc* function of the package {flux}.

3. Results

3.1. Baseline characterization

For the initial data characterization, we worked with the daily time series data. The autocorrelograms showed high autocorrelation at lag 7, 14, 21, etc., indicative of a day-of-week effect. Not surprisingly, the number of declarations was significantly lower on Saturday and Sunday.

At the weekly level, all baselines except TS2 showed a significant positive trend: TS0 had +0.07 declarations per month (p-value = 0.0001), TS1 had +0.06 declarations per month (p-value = 0.01), and TS3 had +0.07 declarations per month (p-value = 0.000). However, this trend was mainly due to the first years of data collection (see Figure 14).

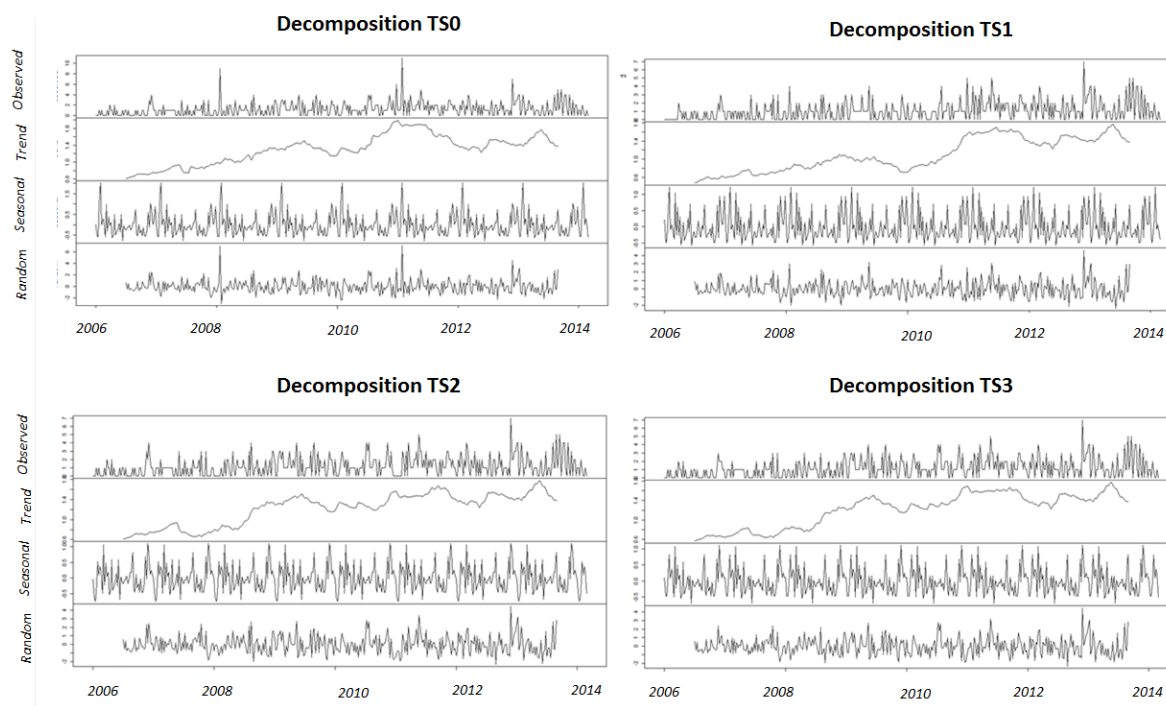


Figure 14 : Decomposition of time series. For each decomposition, the upper graph represents the observed data, the second graph shows the trend identified in the data, the third graph indicates the seasonal pattern found in the data and the fourth graph shows the residuals after deletion of trend and seasonal components.

A significant seasonal effect was also present in all time series: the number of declarations appeared highest in November, December, and January compared to other months. However, this seasonality was weak and principally apparent in the raw TS0 data, due to EHV-1 and EHV-sp outbreaks present in the dataset during the winters of 2008, 2011, and 2013 (see Figure 14).

3.2. Smoothing and forecasting

From 2006 to 2010, the data from each time series were fitted to their respective appropriate regression model, using variables that accounted for seasonal effects. For the Poisson as well as the NB regression, the best fit was obtained for all time series with the simple model:

$$\text{Number_of_cases} \sim \sin(2\pi \cdot \text{week}/53) + \cos(2\pi \cdot \text{week}/53) + \log(\text{histmean})$$

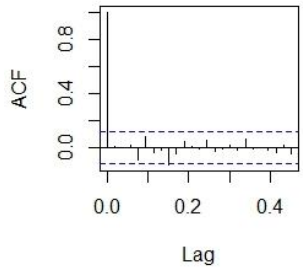
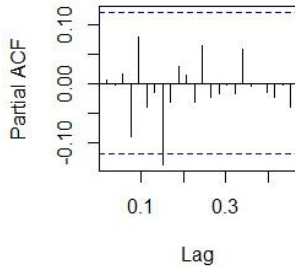
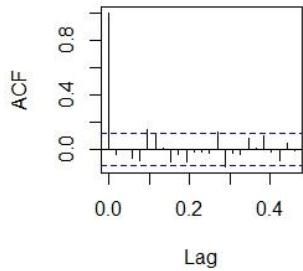
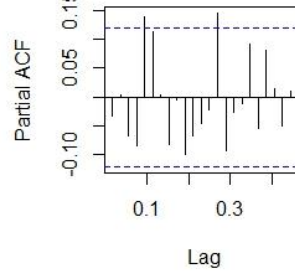
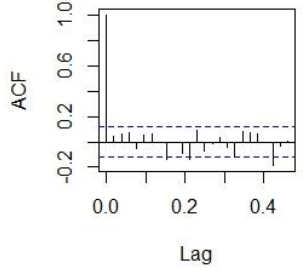
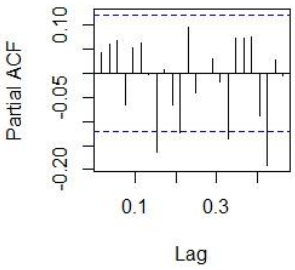
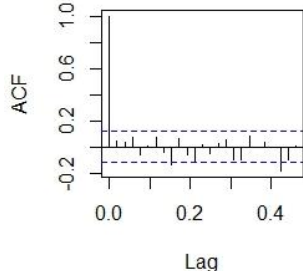
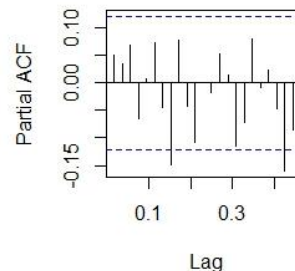
NB and Poisson regressions performed equally well for all time series, with the exception of TS0 (raw data), for which the NB model provided a better fit (AIC 749 vs. 761).

The details of differences between the smoothing performance of the best generalized linear models obtained and HW are presented in Table 10 and Table 11. In all regression methods used, TS0 produced the worst results, while TS1 generated the best fitting parameters. TS2 and TS3 yielded intermediary results, with better performances for TS3 than for TS2.

Table 10 : Smoothing and forecasting performance of GLMs. ACF and PACF are, respectively, the autocorrelation and partial autocorrelation functions of the residuals. Residuals are theoretically assumed to have an ACF and PACF that have no correlation for all lags. RMSE_c and RMSE_v are measures of root-mean-squared error within the calibration period and the validation period, respectively. In both cases, lower values are better.

ACF and PACF				AIC	RMSE _c	RMSE _v
TS0	glm (NB)	ACF of Residuals	PACF of Residuals	735	1.29	1.33
		 The ACF plot shows the autocorrelation of residuals for TS0 using a negative binomial model. The y-axis is labeled 'ACF' and ranges from 0.0 to 0.8. The x-axis is labeled 'Lag' and ranges from 0 to 20. The plot shows a sharp drop to near zero at lag 1, with subsequent lags remaining within the dashed confidence interval lines.	 The PACF plot shows the partial autocorrelation of residuals for TS0 using a negative binomial model. The y-axis is labeled 'Partial ACF' and ranges from -0.10 to 0.10. The x-axis is labeled 'Lag' and ranges from 0 to 20. The plot shows a sharp drop to near zero at lag 1, with subsequent lags remaining within the dashed confidence interval lines.			
TS1	glm (poisson)	ACF of Residuals	PACF of Residuals	635	0.93	1.36
		 The ACF plot shows the autocorrelation of residuals for TS1 using a Poisson model. The y-axis is labeled 'ACF' and ranges from 0.0 to 0.8. The x-axis is labeled 'Lag' and ranges from 0 to 20. The plot shows a sharp drop to near zero at lag 1, with subsequent lags remaining within the dashed confidence interval lines.	 The PACF plot shows the partial autocorrelation of residuals for TS1 using a Poisson model. The y-axis is labeled 'Partial ACF' and ranges from -0.10 to 0.10. The x-axis is labeled 'Lag' and ranges from 0 to 20. The plot shows a sharp drop to near zero at lag 1, with subsequent lags remaining within the dashed confidence interval lines.			
TS2	glm (poisson)	ACF of Residuals	PACF of Residuals	700	1,02	1.34
		 The ACF plot shows the autocorrelation of residuals for TS2 using a Poisson model. The y-axis is labeled 'ACF' and ranges from 0.0 to 0.8. The x-axis is labeled 'Lag' and ranges from 0 to 20. The plot shows a sharp drop to near zero at lag 1, with subsequent lags remaining within the dashed confidence interval lines.	 The PACF plot shows the partial autocorrelation of residuals for TS2 using a Poisson model. The y-axis is labeled 'Partial ACF' and ranges from -0.10 to 0.15. The x-axis is labeled 'Lag' and ranges from 0 to 20. The plot shows a sharp drop to near zero at lag 1, with subsequent lags remaining within the dashed confidence interval lines.			
TS3	glm (poisson)	ACF of Residuals	PACF of Residuals	681	0.96	1.32
		 The ACF plot shows the autocorrelation of residuals for TS3 using a Poisson model. The y-axis is labeled 'ACF' and ranges from 0.0 to 0.8. The x-axis is labeled 'Lag' and ranges from 0 to 20. The plot shows a sharp drop to near zero at lag 1, with subsequent lags remaining within the dashed confidence interval lines.	 The PACF plot shows the partial autocorrelation of residuals for TS3 using a Poisson model. The y-axis is labeled 'Partial ACF' and ranges from -0.10 to 0.15. The x-axis is labeled 'Lag' and ranges from 0 to 20. The plot shows a sharp drop to near zero at lag 1, with subsequent lags remaining within the dashed confidence interval lines.			

Table 11 : Smoothing and forecasting performance of Holt-Winters models. ACF and PACF are, respectively, the autocorrelation and partial autocorrelation functions of the residuals. Residuals are theoretically assumed to have an ACF and PACF that have no correlation for all lags. $RMSE_c$ and $RMSE_v$ are measures of root-mean-squared error within the calibration period and the validation period, respectively. In both cases, lower values are better.

	ACF and PACF		$RMSE_c$	$RMSE_v$
TS0	ACF of Residuals	PACF of Residuals	1.07	1.55
				
TS1	ACF of Residuals	PACF of Residuals	0.77	1.23
				
TS2	ACF of Residuals	PACF of Residuals	0.85	1.32
				
TS3	ACF of Residuals	PACF of Residuals	0.80	1.21
				

3.3. Outbreak detection

The results showed that, for a given method, there was no difference between the different time series tested. However, the generalized linear model always outperformed the Holt-Winters method in terms of detection performance (see details in Table 12).

The AUCs of all methods and time series were low, but it is important to note that the sensitivity used here was based on the number of weeks within an epidemic period that produced an alarm (Se_{wk}). By using instead the percentage of outbreaks detected (with at least one alarm) among all the outbreaks inserted (Se_{out}), the AUCs for all combinations of time series and methods improved to 0.95.

With the generalized linear model, the optimal balance between Se_{wk} , Se_{out} , Sp , and the time-to-detection within an epidemic period was obtained when the alarm threshold exactly equaled the value of the standard error of the model prediction (see details in Table 13). This alarm threshold detected more than 95% of the inserted outbreaks, with an average time to the first true alarm of less than 3 weeks from the start of the outbreak, with the exception of TS0. Specificity varied according to the smoothing method used, but ranged from 0.80 to 0.94 for all generalized linear models. Alarm thresholds that were based on a value higher than the standard error of the model prediction ($K > 1$) had the same detection rate but took more time to produce the first true alarm (> 3 weeks). Instead, an alarm threshold based on a value lower than the standard error of the model prediction ($K < 1$) had the same detection rate and required less time to produce the first true alarm, but resulted in the lowest specificity of all models (from 0.71 to 0.57). The alarm threshold equal to the standard error of the model prediction ($K = 1$) detected from 3 to 6 nervous cases depending on the time series, the smoothing method, and the time period considered.

With the Holt-Winters approach, the optimal balance between Se_{wk} , Se_{out} , Sp , and the time-to-detection in an epidemic period was obtained when the alarm threshold equaled the standard error of the model prediction multiplied by a constant of 0.5 (see details in Table 14). This alarm threshold detected more than 95% of the inserted outbreaks and the average time-to-detection was less than 3 weeks from the start of the outbreak. The associated specificity had an average value of 0.87. Alarm thresholds that were based on constants higher than 0.5 had the same detection rate but needed more time to produce the first true alarm (> 3 weeks).

Table 12 : System's detection performances. Se_wk = sensitivity based on detection of each week which is a part of an epidemic period, Sp = specificity based on the number of true negative and false positive alarms, ROC = receiver operating characteristic, AUC = Area Under the receiver operating characteristic curve.

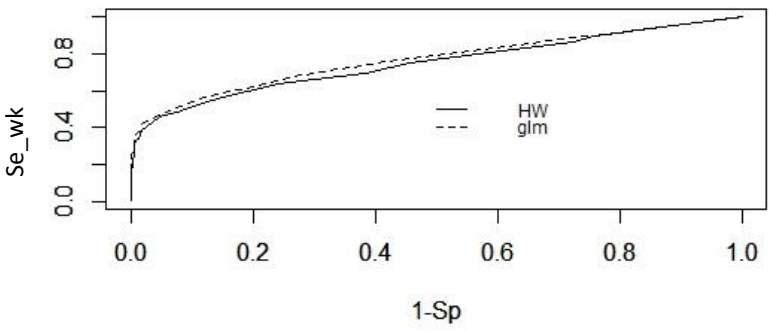
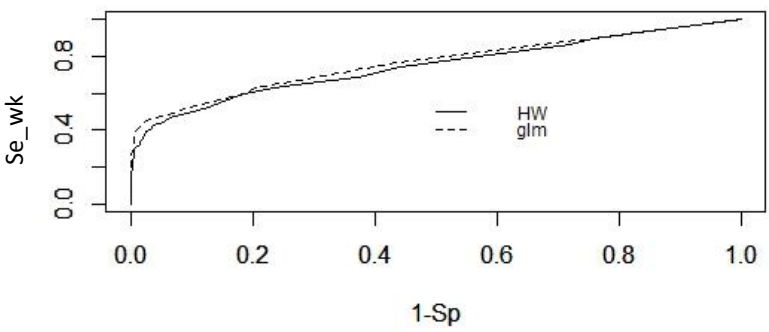
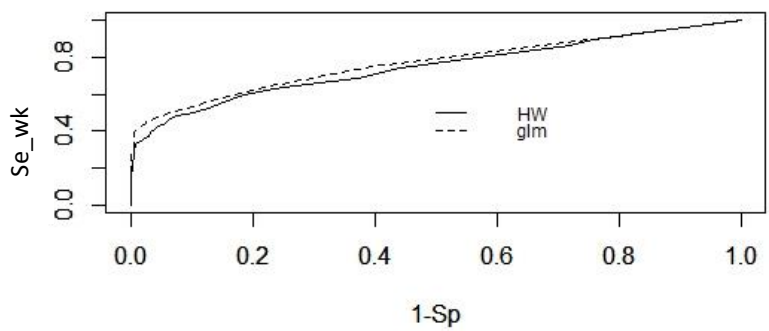
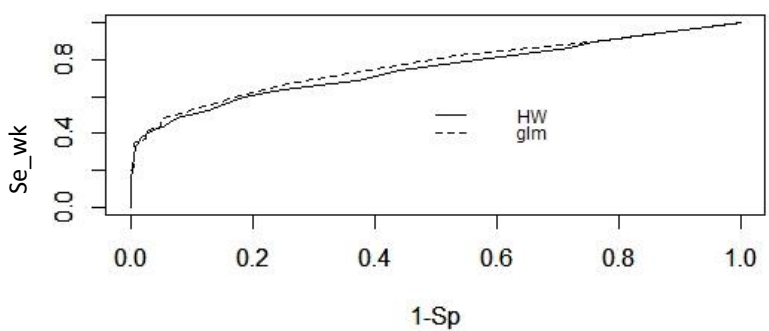
AUC		ROC curves	
TS0	Glm(NB) 0.765617		
	HW 0.747400		
TS1	Glm(poisson) 0.763103		
	HW 0.7452638		
TS2	Glm(poisson) 0.764747		
	HW 0.7450518		
TS3	Glm(poisson) 0.7632676		
	HW 0.745603		

Table 13 : system's detection performances for time series fitted with generalized linear models. K = constant multiple of the standard error of the model prediction, Se_wk = sensitivity based on detection of each week which is a part of an epidemic period, Se_out = sensitivity based on the number of outbreaks detected, Sp = specificity based on the number of true negative and false positive alarms, Average time of detection = Average number of weeks needed to produce the first alarm within an outbreak.

K		Se_wk	Se_out	Sp	Average time of detection (weeks)
0.5	TS0	0.68	0.96	0.71	1.83
	TS1	0.76	0.96	0.57	1.56
	TS2	0.76	0.96	0.57	1.57
	TS3	0.76	0.96	0.57	1.56
1	TS0	0.48	0.96	0.94	3.13
	TS1	0.60	0.96	0.80	2.36
	TS2	0.62	0.96	0.80	2.16
	TS3	0.56	0.96	0.86	2.75
1.5	TS0	0.33	0.96	0.99	3.68
	TS1	0.44	0.96	0.98	3.27
	TS2	0.44	0.96	0.97	3.25
	TS3	0.41	0.96	0.97	3.33
2	TS0	0.22	0.88	1	4.10
	TS1	0.37	0.96	0.99	3.38
	TS2	0.39	0.96	0.99	3.34
	TS3	0.31	0.94	0.99	3.73
2.5	TS0	0.11	0.74	1	4.72
	TS1	0.26	0.92	1	3.85
	TS2	0.27	0.93	1	3.77
	TS3	0.20	0.84	1	4.21
3	TS0	0.07	0.60	1	4.38
	TS1	0.21	0.89	1	4.23
	TS2	0.22	0.89	1	4.01
	TS3	0.15	0.80	1	4.69

Table 14: system's detection performances for time series fitted with Holt-Winters. K = constant multiple of the standard error of the model prediction, Se_wk = sensitivity based on detection of each week which is a part of an epidemic period, Se_out = sensitivity based on the number of outbreaks detected, Sp = specificity based on the number of true negative and false positive alarms, Average time of detection = Average number of weeks needed to produce the first alarm within an outbreak.

K		Se_wk	Se_out	Sp	Average time of detection (weeks)
0.5	TS0	0.53	0.96	0.88	2.79
	TS1	0.52	0.96	0.87	2.73
	TS2	0.52	0.96	0.87	2.73
	TS3	0.52	0.96	0.87	2.73
1	TS0	0.33	0.96	0.98	3.57
	TS1	0.33	0.94	0.98	3.52
	TS2	0.33	0.94	0.99	3.55
	TS3	0.33	0.94	0.99	3.58
1.5	TS0	0.18	0.86	0.99	4.27
	TS1	0.17	0.84	0.99	4.39
	TS2	0.17	0.84	0.99	4.39
	TS3	0.16	0.84	0.99	4.39
2	TS0	0.09	0.64	0.99	4.45
	TS1	0.09	0.65	0.99	4.63
	TS2	0.09	0.65	0.99	4.63
	TS3	0.09	0.65	0.99	4.66
2.5	TS0	0.06	0.50	1	4.36
	TS1	0.03	0.32	1	4.19
	TS2	0.05	0.48	0.99	4.14
	TS3	0.05	0.48	0.99	4.17
3	TS0	0.03	0.37	1	4.09
	TS1	0.017	0.18	1	4.50
	TS2	0.03	0.32	1	4.19
	TS3	0.03	0.32	1	4.19

4. Discussion

As expected, the preprocessing methods that were used to remove past outbreaks present in the dataset modified the seasonality of the time series. Indeed, outbreaks of EHV-1 that were present in TS0 were mainly reported during winter, which is consistent with reports of seasonal patterns of disease outbreaks from a recent consensus statement (Lunn et al. 2009). Removing these outbreaks from the TS0 data decreased the impact of season on the baseline and improved the smoothing performance of the two forecasting methods tested. The raw data (TS0) produced the worst results compared to all time series in which outbreaks were removed. Regarding the impact of preprocessing methods on system performance, no impact was observed when Holt-Winters smoothing was used. Conversely, the detection performance differed between time series fitted with glm: TS0 obtains always the worst results (longest time-to-detection and lowest Se_wk/Sp).

RESPE currently uses raw data (TS0) and an alarm threshold of four declarations per week for the detection of outbreaks. According to the analysis performed in this study with TS0, this alarm threshold is close to the standard error of our model prediction using GLMs (between four and five cases, depending on season). The current detection performance of the RESPE system is thus: weekly sensitivity close to 0.48, specificity close to 0.94, and average time-to-detection close to 3.13 weeks. These values are lower than those obtained here with preprocessed data (TS1, TS2, TS3) and reveal the importance of data preprocessing in improving outbreak detection for syndromic surveillance.

When we use values of Se_{wk} to evaluate overall system's performance, it is clear that the syndromic surveillance system suffers from low sensitivity. This is not surprising as, during a WNV outbreak, the number of cases observed per week of the outbreak is generally low, especially during the initial and final stages of the outbreak (see figure 12). When considering the system's ability to detect an outbreak as a single entity (Se_{out}), all combinations of time series and methods used were able to detect more than 95% of the inserted outbreaks. However, compared to HW, GLM enabled earlier outbreak detection, with a better Se_{wk} , for a given Sp . In the end, the best performances were obtained using GLMs associated with preprocessing methods TS1, TS2, or TS3, and an alarm threshold set to the standard error of the model prediction. With these settings, the surveillance system can detect 96% of outbreaks, with an average time-to-detection of 2.16 to 2.75 weeks, weekly sensitivity (Se_{wk}) between 0.56 and 0.62, and specificity ranging from 0.80 to 0.86. A K value (multiple of the standard error) between 2 and 3.5 is often chosen to ensure a false alarm rate below 5% (Mandl et al. 2004a) but, in our case, this threshold increased the time-to-detection to between 3 and 5 weeks, an undesirable outcome when striving for early detection. In this study, we did not determine which time series (TS1, TS2, or TS3) was the most efficient, as such a decision would be made in real life by decision makers and would depend on the objectives of the surveillance.

Our study shows that the RESPE data on nervous symptoms in horses can be used as an alarm system for WNV outbreaks in France and the full assessment of system's performance was possible thanks to simulated data. This is the first time that a real assessment of system performance has been implemented for WNV surveillance. Previous early warning systems developed for WNV only identified risk factors of WNV outbreaks, but did not evaluate the detection performances of those systems (Adlouni et al. 2007; Bellini et al. 2014a; Brown 2012; Chaskopoulou et al. 2013; Gosselin et al. 2005; Rosà et al. 2014; Shuai et al. 2006; Valiakos et al. 2014). Timeliness has occasionally been evaluated but only based on a limited number of real WNV outbreaks, and has not been associated with a further assessment of system performance (Calzolari et al. 2013; Chaintoutis et al. 2014; Eidson et al. 2001; Johnson et al. 2006; Mostashari et al. 2003; Veksler et al. 2009). Only one attempt (Leblond et al. 2007) to assess the sensitivity and specificity of surveillance has been made but the parameters of interest were only evaluated based on a limited number of outbreaks, which did not allow any conclusions to be drawn regarding overall system performance. The outbreaks were simulated using real data and should thus be consistent with the course of a real WNV outbreak. However, there is always the risk that the conditions used here to evaluate

system performance were unrealistic, and thus other outbreaks, differing in magnitude and form, should be also tested in order to confirm our results.

5. Conclusion

A classical syndromic surveillance approach based on nervous symptoms in horses can be implemented using RESPE data in order to detect WNV outbreaks. The results produced in our study are better than those obtained with the current detection system and argue in favor of data preprocessing for the improvement of outbreak detection.

The output of the syndromic system produced here was a yes/no qualitative output: “No, there is no outbreak” or “Yes, something unusual is happening in the population”. This output has the advantage of simplicity, but its interpretation/utilization may be complex when data are close to the alarm threshold. This output also has a low specificity. Indeed, we can detect WNV but probably also equine herpesvirus. It would be interesting to combine the output from syndromic surveillance with other information to reach a better specificity. However, it is currently complicated to combine this output with other epidemiological knowledge, such as disease seasonality or environmental risk factors, parameters that are fundamental when working with vector-borne diseases like WNV. This dilemma will be addressed in the next section.

C. VALUE OF EVIDENCE

1. Introduction

To address the problems associated with the qualitative outputs of syndromic surveillance and the difficulties of combining syndromic surveillance with other epidemiological knowledge, we tested a new approach for outbreak detection based on Bayes' rule. Bayes' rule is already used in a wide range of disciplines to combine information and provide numerical estimation of a likelihood ratio. In the framework of Bayesian analyses, this likelihood ratio is used to test hypotheses and clearly specify the strength of forensic evidence for/against a hypothesis.

In the next section, we evaluate the applicability of the Bayesian likelihood ratio framework to the detection of outbreaks in syndromic surveillance. Two examples are considered: nervous syndromes in horses as an early warning of WNV outbreaks, and respiratory syndromes in horses for the detection of equine influenza. The objective was, first, to build a more objective, flexible, and easily interpretable output for syndromic surveillance and, second, to combine syndromic surveillance with other epidemiological knowledge.

This part of the work was implemented in collaboration with Gunnar Andersson (SVA, Sweden) and published in PLOSOne (see below). This work also resulted in one oral presentation at the annual conference of AEEMA ("Association pour l'étude de l'épidémiologie des maladies animales", a Francophone association dedicated to veterinary epidemiology) and one publication in their journal, '*Bulletin Epidémiologie et Santé Animale*' (text available in French in Appendix 10).

Using Bayes' Rule to Define the Value of Evidence from Syndromic Surveillance

Mats Gunnar Andersson^{1*}, Céline Faverjon², Flavie Vial³, Loïc Legrand^{4,5}, Agnès Leblond^{5,6}

1 Department of Chemistry, Environment and Feed Hygiene, The National Veterinary Institute, Uppsala, Sweden, **2** INRA UR346 Animal Epidemiology, VetagroSup, Marcy L'Etoile, France, **3** Veterinary Public Health Institute, DCR-VPH, Vetsuisse Fakultät, Bern, Switzerland, **4** LABÉO - Frank Duncombe, Unité Risques Microbiens (U2RM), EA 4655, Normandie Université, Caen, Normandy, France, **5** Réseau d'Epidémiologie en Pathologie Equine (RESPÉ), Caen, France, **6** INRA UR346 Animal Epidemiology et Département Hippique, VetAgroSup, Marcy L'Etoile, France

Abstract

In this work we propose the adoption of a statistical framework used in the evaluation of forensic evidence as a tool for evaluating and presenting circumstantial "evidence" of a disease outbreak from syndromic surveillance. The basic idea is to exploit the predicted distributions of reported cases to calculate the ratio of the likelihood of observing n cases given an ongoing outbreak over the likelihood of observing n cases given no outbreak. The likelihood ratio defines the Value of Evidence (V). Using Bayes' rule, the prior odds for an ongoing outbreak are multiplied by V to obtain the posterior odds. This approach was applied to time series on the number of horses showing clinical respiratory symptoms or neurological symptoms. The separation between prior beliefs about the probability of an outbreak and the strength of evidence from syndromic surveillance offers a transparent reasoning process suitable for supporting decision makers. The value of evidence can be translated into a verbal statement, as often done in forensics or used for the production of risk maps. Furthermore, a Bayesian approach offers seamless integration of data from syndromic surveillance with results from predictive modeling and with information from other sources such as disease introduction risk assessments.

Citation: Andersson MG, Faverjon C, Vial F, Legrand L, Leblond A (2014) Using Bayes' Rule to Define the Value of Evidence from Syndromic Surveillance. PLoS ONE 9(11): e111335. doi:10.1371/journal.pone.0111335

Editor: Simon Gubbins, The Pirbright Institute, United Kingdom

Received: June 11, 2014; **Accepted:** September 23, 2014; **Published:** November 3, 2014

Copyright: © 2014 Andersson et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. All relevant data are within the paper and its Supporting Information files.

Funding: This work was funded by EMIDA ERA-NET through the Swedish research council Formas (<http://www.formas.se/en/>) (Contract no. 221-2011-2214) to MGA; EMIDA ERA-NET through Netherlands Food Consumer Product Safety Authority (www.nvwa.nl/english) and the Dutch Ministry of Economic Affairs (<http://www.government.nl/ministries/ez>) (project no. BO 20-009-009), and IFCE (Institut Français du Cheval et de l'Équitation) (<http://www.ifce.fr/>) to CF; INRA (Animal epidemiology Unit – EPIA) (<http://www.inra.fr/en/Scientists-Students>) and IFCE (Institut Français du Cheval et de l'Équitation) (<http://www.ifce.fr/>) to AL; 'Conseils Généraux du Calvados, de la Manche et de l'Orne' (County Councils) (<http://www.calvados.fr/cms>) to LL; Veterinary Public Health Institute, DCRVPH, Vetsuisse Fakultät, (http://www.vetsuisse.unibe.ch/vphi/content/index_ger.html) to FV; Labéo Frank Duncombe (<http://www.labo-frank-duncombe.fr/>) provided support in the form of a salary for author LL. The funders did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the 'author contributions' section. The specific roles of these authors are articulated in the 'author contributions' section.

Competing Interests: LL is an employee of LABÉO. LABÉO was funded by the Conseils Généraux du Calvados, de la Manche et de l'Orne (County Councils). There are no patents, products in development or marketed products to declare. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

* Email: gunnar.andersson@sva.se

These authors contributed equally to this work.

Introduction

Syndromic surveillance appeared in the late 1990's and is becoming more and more popular in a wide range of human public health issues such as seasonal disease surveillance [1] and digital disease surveillance [2]. The wider acceptance of the relevance of the "One Health" concept [3] amongst public health practitioners has led to an increased exchange of methodologies and disease control knowledge between the human medicine and the veterinary sides. In the last 5 years, researchers in veterinary medicine have been investigating the application of syndromic surveillance methods for the early detection of zoonotic and non-zoonotic diseases [4].

There is no unique definition of "syndromic surveillance" but it is commonly accepted that it focuses on data collected prior to clinical diagnosis or laboratory confirmation [5,6]. It is therefore based on non-specific health indicators which result in a

surveillance system with low specificity but allow the early detection of outbreaks without *a priori* considerations. This constitutes a major advantage over traditional approaches which focus on a disease, or a list of reportable diseases, and rely on the ability of clinicians to correctly diagnose cases, which may be difficult when faced with a rare or emerging disease [4]. Moreover, the systematic and continuous data collection and analysis processes reduce the impact of chronic under-reporting observed in classical passive surveillance systems and also increases the sensitivity of this method [4]. Syndromic surveillance does not replace traditional approaches to disease monitoring (e.g. risk-based, active etc...) but is seen as an interesting and complementary tool for outbreak detection with a low specificity but with better sensitivity and timeliness [7].

Current approaches used in syndromic surveillance first seek to define the normal properties of the syndrome time-series when no

outbreak of disease is recorded [4,6] in order to be able to detect abnormal events overlaid on top of the background noise during an outbreak situation. In traditional aberration detection methods, an alarm goes off when the observed data exceed the expected values from the population [4,6]. Such algorithms have an epidemic threshold and provide a yes/no qualitative output: “No, there is no outbreak” or “Yes, something unusual is happening in the population”.

This black or white vision of the health of the population of concern is simple but it may not always be adequate or useful for decision makers who may often find themselves in grey areas (indicator values close to the epidemic threshold). Moreover, binary result can also be difficult to combine with other epidemiological knowledge such as a probability of disease introduction or other complex parameters which influence decision making [8]. The development of syndromic surveillance quantitative outputs, which are more objective, flexible and easily interpretable, is a promising area of research.

The art of presenting scientific evidence to decision makers has been more extensively studied in forensic sciences in which legal certainty requires statements that clearly specify how strong the evidence for/against an hypothesis is and how the expert reached that conclusion. In recent years, the state of the art in forensic interpretation has been to evaluate forensic evidence using likelihood ratios in the framework of Bayesian hypothesis testing. Within this framework, it evaluates the extent to which results from forensic investigations speak in favor of the prosecutors or defendants hypotheses [9,10]. The Bayesian approach has been applied to a wide range of forensic problems including evidence based on DNA analysis [10], mass spectroscopy [11], transfer of glass, fibers and paint [10] and microbial counts [12]. However, although initially developed for the legal system, the approach has been identified as useful for supporting decision making in other situations such as the tracing of *Salmonella spp* [13].

The aim of this study is to test the applicability of the Bayesian likelihood ratio framework to the early detection of outbreaks in a syndromic surveillance system. Transferability of the method is demonstrated by using two examples based on real data coming from RESPE, the French surveillance network on equine diseases. The first example makes use of data on French horses presenting nervous symptoms (NeurSy) and aim to test the ability of our approach to detect simulated outbreaks of an exotic disease, West Nile Virus (WNV). West Nile disease is an important zoonotic disease and syndromic surveillance applied in horses could be used as an early warning system to protect the human population [14]. The second example focuses on data on French horses with respiratory symptoms (RespSy) and is used to detect outbreaks of divergent strains of equine influenza (New-Influenza), a non-zoonotic disease leading to vaccine failure [15–18].

Materials and Methods

Background theory and proposed framework

Forensic evaluation of evidence is based on Bayesian hypothesis testing. In a syndromic surveillance context, this would mean that, in a particular week, there are two mutually exclusive hypotheses that should be evaluated, for example: H_1 “There is an ongoing outbreak of disease x” and H_0 “There is NOT an ongoing outbreak of disease x”. Without any extra information, the relative probability of the two hypotheses may be expressed as the *a priori* odds:

$$O_{pri} = \frac{P(H_1)}{P(H_0)} \quad (\text{Eq.1})$$

where

$P(H_1)$: The *a priori* probability for hypothesis H_1 . Typically the probability of an ongoing outbreak of the disease of interest in a particular region.

$P(H_0)$: The *a priori* probability for hypothesis H_0 which is the complementary hypothesis to H_1 . Typically the probability of an outbreak NOT going on.

In other words, the *a priori odds* define our *prior belief* about the disease status in the region. In a typical situation, the prior odds would be low (e.g. 1:1000) but under some circumstances, it might be higher (e.g. if an outbreak is ongoing in a neighboring country). When we are presented evidence (E) of some kind pointing in favor (or against) of H_1 , this will make us update our belief. This posterior belief is expressed as the *a posteriori* odds.

$$O_{post} = \frac{P(H_1|E)}{P(H_0|E)} \quad (\text{Eq.2})$$

Where:

$P(H_1|E)$ is the probability of hypothesis H_1 , given the evidence (E).

$P(H_0|E)$ is the probability of hypothesis H_0 , given the evidence (E).

In syndromic surveillance, the evidence (E) is typically the number of reported suspected cases in a given time period. The degree to which the posterior belief differs from the prior belief will depend on the strength of the evidence. If the evidence is weak, the posterior odds will be similar to the prior odds whereas strong evidence in favor of H_1 would result in posterior odds being much higher than the prior odds. At this point, it is important to note that the hypotheses to evaluate (H_1) may differ and that the interpretation of the same piece of evidence would depend on the choice of H_1 . For example 10 reported cases of syndromes in horses may be a strong evidence that there is something unusual going on if these are nervous cases (H_1 = “ongoing outbreak of some nervous disease (i.e. WNV)”) but only weak evidence in favor of an equine influenza in the case of a respiratory syndrome (H_1 = “ongoing outbreak of equine influenza”), since in the latter case we might have expected far more reported cases.

This intuitive reasoning can be formalized by the application of Bayes’ theorem:

$$O_{post} = V \times O_{pri} = \frac{P(H_1|E)}{P(H_0|E)} = \frac{P(E|H_1)}{P(E|H_0)} \times \frac{P(H_1)}{P(H_0)} \quad (\text{Eq.3})$$

Where:

E is the number of reported cases of a syndrome in the particular week.

$P(E|H_1)$ is the probability of observing the evidence (E) given that H_1 is true.

$P(E|H_0)$ is the probability of observing the evidence (E) given that H_0 is true

In order to estimate $P(E|H_1)$ and $P(E|H_0)$ we need information on the probability distribution for the number of reported cases in a non-outbreak and outbreak situation. The probability of observing n cases given that H_1 is true can be estimated using statistical modeling of baseline data [19]. When the cases are

independent (i.e. not clustered), the data can be modeled using a general dynamic Poisson model [19]. When cases are clustered (overdispersion), the Poisson model will underestimate the probability of observing very high or very low number of cases, and in such cases, the data can be modeled by continuous mixtures of the Poisson distribution including Negative Binomial (NB) distribution or Poisson-log-normal (PLN) distribution [19].

The probability of E (observation of n cases) during an outbreak is calculated as:

$$P(E|H_1) = \sum_{i=0}^n P_{base}(i) \times P_{out}(n-i) \quad (\text{Eq.4})$$

Where

$P_{base}(i)$ = Probability of drawing i cases from the baseline distribution (e.g. Poisson(λ) or NB($\mu = \mu_{base}$, size = θ_{base}))

$P_{out}(i)$ = Probability of drawing i cases from the outbreak distribution (e.g. NB($\mu = \mu_{out}$, size = θ_{out}))

The outbreak distribution may be estimated by fitting an appropriate probability distribution to data from historical outbreaks. In the absence of data, the outbreak distribution may be defined based on expert knowledge about the disease in question or assumptions about the distribution of a new disease. In most cases there would be a large uncertainty about the shape of the outbreak distribution.

The next estimate is the probability of observing the Evidence (E) that is the actual number of reported cases. In forensics, the value of evidence (V) is defined as the ratio between the posterior and prior odds for H_1 versus H_0 . The value of evidence (Fig. 1, line Log(V)) can be calculated from the two distributions by dividing the probabilities for each number of observed cases using equation 5:

$$V = \frac{P(E|H_1)}{P(E|H_0)} \quad (\text{Eq.5})$$

As illustrated in Fig 1 the value of evidence will depend on the assumptions about the outbreak. In the examples A to D, 10 cases are reported from a region where the baseline prevalence is around 5 cases per week. If it is expected that an outbreak may be small, resulting in only a small number of extra cases, 10 reported cases would speak in favor of an outbreak (Fig. 1, A, C). If, on the other hand, the disease(s) of interest are expected to yield a relatively large number of cases the evidence would speak against an outbreak (Fig. 1, B, D).

In addition, the strength of the evidence will depend on the precision on the estimates for the number of outbreak-related cases. If the distributions are wide (low θ , Fig 1A, 1B), the absolute value of $\log(V)$ is smaller whereas more narrow distributions (high θ , Fig 1C, 1D) result in higher values of $\log(V)$. This is intuitive: the more we know about what we expect to see during an outbreak, the stronger conclusions we will make from the observed evidence.

Using the value of evidence for decision making

In contrast to traditional outbreak detection algorithms, the value of evidence approach does not have a built-in decision threshold. Typically a decision maker would not act upon syndromic surveillance data alone but rather combine it with other available knowledge. Cameron [20] proposed several approaches to disease freedom questions: (1) population or surveillance sensitivity, (2)

probability of freedom from disease, and (3) expected cost of error – i.e., consequences of false positive and false negative results. All approaches underline how the value of inspection findings will be augmented when interpreted in a broader context to complement other monitoring and surveillance systems (MOSS) activities. One option for a decision maker would be to set an action threshold for the posterior odds. We might, for example, want to initiate an epidemiological investigation if the odds that there is an ongoing outbreak are larger than 1:1 or 1:100. Ideally the decision maker would make a cost-benefit analysis taking into account the expected costs for taking action versus not taking action. For example the decision maker may initiate control measures (vaccination program *etc*) when the odds are such that, on average, the reduced loss from the early detection of the outbreak would exceed the extra costs from initiating control measures (or vaccination programs) in response to false alarms.

The combination of evidence evaluation and decision theory is discussed in [21]. The expected utility (\bar{u}) of action a_i is the average amount of loss that we expect to incur with this action. In the context of diseases surveillance, an action could be to implement movement restrictions, vaccination, sampling, control of vectors or to do nothing. The loss could be the direct financial losses (e.g. animal infection, disease and production losses) but also the indirect losses (e.g. surveillance and control costs, compensation costs, potential trade losses, social consequences). Since an unmanaged outbreak as well as actions will result in costs, the expected utility will always be zero or negative. In this framework the expected utility (\bar{u}) of action a_i is defined as:

$$\bar{u}(a_i|\cdot) = \sum_{j=0}^1 u(C_{ij})p(H_j|\cdot) \quad (\text{Eq.6})$$

where

H_1 = Outbreak

H_0 = No outbreak

a_0 = No action

a_1 = Action

C_{ij} = Different scenarios with respect to hypothesis on outbreak status (H_0, H_1) and action (a_0, a_1) C_{00} represents the case with no disease and no action implemented. C_{01} is no disease but action implemented, C_{10} disease but no action and C_{11} is disease and action implemented)

$p(H_j|\cdot)$ = probability of hypothesis j given all available knowledge (Prior probability & evidence)

$u(C_{ij})$ = expected utility for each possible situation C_{ij} . Since gain is zero the utility is determined by economical and socio-economical loss.

According to this framework it is favorable to act when the expected utility of action ($\bar{u}(a_1|\cdot)$) is higher than the expected utility of no action ($\bar{u}(a_0|\cdot)$). The relation between posterior probability ($P(H_1|E)$) and posterior odds (O_{post}) is defined by:

$$O_{post} = \frac{P(H_1|E)}{1 - P(H_1|E)} \quad (\text{Eq.7})$$

and

$$P(H_1|E) = \frac{O_{post}}{1 + O_{post}} \quad (\text{Eq.8})$$

Thus equation 6 can be reformulated as

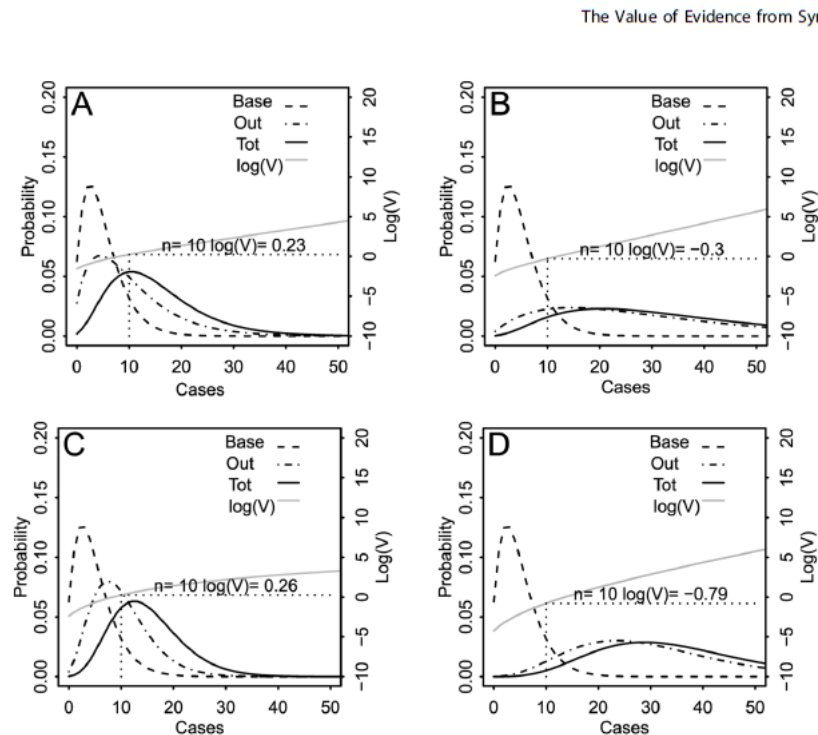


Figure 1. Value of evidence (V) and probability of observing 10 cases during a non-outbreak (Base) and outbreak situation (Out) with different assumptions about the magnitude of an outbreak. The baseline cases are distributed according to NB $\mu = 5$, $\theta = 2.55$. The value of evidence, $\log(V)$ is calculated as $\log_{10}(p(n|outbreak)/p(n|baseline))$. The distribution during an outbreak (Tot) is the sum of baseline cases and outbreak cases. In the examples A to D outbreak related cases are distributed according to (A) NB($\mu = 10$, $\theta = 2$), (B) NB($\mu = 30$, $\theta = 2$), (C) NB($\mu = 10$, $\theta = 5$), (D) NB($\mu = 30$, $\theta = 5$). doi:10.1371/journal.pone.0111335.g001

$$\bar{u}(a_i|\cdot) = \sum_{j=0}^1 u(C_{ij}) \times \frac{O_{post}}{(1 + O_{post})} \quad (\text{Eq.9})$$

For each value of O_{post} the expected utility for action a_1 and a_0 is defined by eq. 9. The expected loss for each situation C_{ij} is based on expert opinion as indicated in table 1. An action threshold for posterior odds (O_{post}^*) can be defined as the value of O_{post} where

$$\bar{u}(a_1|\cdot) = \bar{u}(a_0|\cdot)$$

In this work O_{post}^* was determined by numerical optimization. The derived action threshold for the value of evidence V^* is calculated as:

$$V^* = \frac{O_{post}^*}{O_{pri}} \quad (\text{Eq.10})$$

where the prior odds for an ongoing outbreak $\log_{10}(O_{pri})$ is based on historical experience as well as knowledge about risk factors.

To make a decision, the risk manager would multiply the prior odds with the value of evidence using eq.3 to obtain the posterior odds for an outbreak $O_{post}(H_1|E)$. If this odds goes over the action

threshold $\log_{10}(O_{post}^*)$ where the expected utility from acting exceeds the utility for not acting, a decision would be taken to act.

Performance assessment

Sensitivity, specificity and predictive values of positive and negative tests are important concepts when planning animal health monitoring. In the syndromic surveillance context a true positive (TP) is when the system alerts when an outbreak is ongoing. A true negative (TN) is no alert and no outbreak. A false negative (FN) is when the system does not alert when an outbreak is ongoing, and, false positive (FP) is when the system alerts in the absence of an outbreak.

Sensitivity (SE) is the probability that a true outbreak triggers an alert:

$$SE = TP / (TP + FN) \quad (\text{Eq.11})$$

Specificity (SP) is the probability the there is no alert when no outbreak is ongoing:

$$SP = TN / (TN + FP) \quad (\text{Eq.12})$$

The positive predictive value (PPV) is the probability of an indicated outbreak being a true outbreak:

Table 1. Expected utility associated with different actions and the derived decision threshold & decision.

	Scenario A		Scenario B		Scenario C Large	
	Small outbreak in Autumn		Medium outbreak in Winter		outbreak in Spring	
$u(C_{00}) \text{ Out- act-}$	0		0		0	
$u(C_{10}) \text{ Out- act+}$	-0.5 M€		-0.5 M€		-0.5 M€	
$u(C_{01}) \text{ Out+ act-}$	-5.1 M€		-5.3 M€		-10.1 M€	
$u(C_{11}) \text{ Out+ act+}$	-3.9 M€		-4.1 M€		-6.3 M€	
Action threshold $\text{Log}_{10}(O_{\text{post}}^*)$	-0.38		-0.38		-0.88	
$\text{Log}_{10}(O_{\text{pri}})$	-0.99		-3.03		-1.78	
Action Threshold $\text{Log}_{10}(V^*)$	0.61		2.65		0.9	
Weeks	w36	w39	w1	w4	w25	w28
Cases observed per week	3	4	5	7	5	7
$\text{Log}_{10}(V)$	0.23	0.67	1.30	2.77	1.77	3.41
$\text{Log}_{10}(O_{\text{post}})$	-0.76	-0.34	-1.71	-0.34	-0.01	1.63
Action? $V > V^*$	No	Yes	No	Yes	Yes	Yes

doi:10.1371/journal.pone.0111335.t001

$$PPV = \frac{TP}{(TP + FP)} \quad (\text{Eq. 13})$$

The Negative predictive value (NPV) is the probability that no signal of outbreak is true absence of an outbreak:

$$NPV = \frac{TN}{(TN + FN)} \quad (\text{Eq. 14})$$

The PPV and NPV depend on the (prior) probability of an outbreak and in the performance assessment PPV was calculated as:

$$PPV = \frac{P_{\text{pri}} \times SE}{(1 - P_{\text{pri}}) \times (1 - SP) + SE \times P_{\text{pri}}} \quad (\text{Eq. 15})$$

where:

P_{pri} = prior probability of ongoing outbreak in the week of interest

Implementation

Models were implemented in R x64 version 3.0.2 [22]. The R-Scripts are included as part of the material (Script S1, S2, S3, S4).

Dynamic regression was performed with function *glm* (package {stats}) [22] for Poisson regression and *glm.nb* (package {MASS}) [23]. The expected number of counts at time x were estimated with the *predict* function of the respective package. Alternative regression models were evaluated using the Akaike information criterion (AIC). In addition adjusted deviance (Deviance/df) was used as a measure of goodness of fit (GOF).

The receiver operating characteristic (ROC) curve was generated in R by simulation. Counts for negative weeks were sampled from a Poisson distribution (function *rpois* in package {stats}) with lambda equal to the predicted value for each week in 2011 and 2012 ($n=53000$). Counts for positive weeks were generated by sampling values from the fitted outbreak distribution (function *rbinom* in package {stats}) and adding to the baseline.

SE and SP were calculated for values of $\text{Log}_{10}(V)$ between -1 and +3 in steps of 0.01. The expected PPV for each value of V was calculated as above using the prior odds for outbreak from three scenarios.

Threshold values for posterior odds (O_{post}^*) were estimated using the Solver function of Microsoft Excel 2007.

Sources of data

As a proof of principle the value of evidence framework was applied to neurological and respiratory syndromes in French horses. The associated time series are named NeurSy and RespSy, respectively. These data are collected through the passive surveillance system "RESPE", the French network for the surveillance of equine diseases (<http://www.respe.net/>). This system collects the declarations from veterinary practitioners registered as sentinels who fill online a standardized questionnaire depending on the syndrome concerned. Along with their declaration, veterinarians send standardized samples for the laboratory diagnosis. Tests for equine influenza, equine herpes 1 and 4 and equine arteritis viruses are implemented in the case of a respiratory syndrome, West Nile and equine herpes 1 viruses in the case of a nervous syndrome. In our study, we used these weekly time series.

Data from 2006 to 2010 were used to train our models and define the background noise of each time series when no outbreak occurs. We only used the data on the number of cases with no positive laboratory test result in order to remove the outbreaks from our datasets and obtain these outbreak free baselines. Then, different regression models were tested.

No real outbreak of West Nile disease and divergent strains of equine influenza (New-Influenza) occurred during this time. Instead fictive test data were used for demonstrating outbreak detection. The baselines in the test data were based on NeurSy and RespSy data from 2011 to 2012 where unexplained aberrations, not related to the diseases of interest, were filtered out and fictive outbreaks inserted based on historical data. The weekly counts from several real outbreaks were fitted together to model the outbreaks of each disease. The *prior odds* for each example are based on our knowledge on the epidemiology and risk factors for transmission of the disease. New-Influenza is supposed to have the same probability of occurrence over the year and the

prior odds is thus considered as constant over time. West Nile disease transmission is linked to the vector's level of activity and is thus a seasonal disease. Different *prior odds* are set for each season for this disease.

Data Accessibility

The datasets supporting this article have been uploaded as part of the Material. The baseline data for NeurSy and RespSy are included in Table S1 and Table S2 respectively. The outbreak data for NeurSy and RespSy are included in Table S3 and Table S4 respectively.

The software R can be freely downloaded from the CRAN homepage (<http://cran.r-project.org/>).

Results

Case study – Neurological syndromes and WNV (NeurSy)

Non-outbreak situation. To define the background noise of the NeurSy time series when no outbreak occurred, we fitted alternative regression models based on Poisson and NB distributions from years 2006–2010 on data containing only cases with no positive laboratory results (figure S1). The models evaluated including sinod models with 1, 2 and 3 periods/year and season or month as factorial variables. To account for differences between years we dynamically calculate the average counts for 53 consecutive weeks (*histmean*). To ensure that an ongoing outbreak will not influence the estimate, we used a 10 week guard band [24] for calculation of *histmean*. For the Poisson as well as the NB regression the best fit were obtained with the simplest model:

$$\text{counts} \sim \sin(2\pi t) + \cos(2\pi t) + \log(\text{histmean})$$

where t is time in years. For the Poisson regression we obtained: AIC = 637.8, GOF(adjusted dev) = 1.156. For NB regression the corresponding parameters were: AIC: 639; GOF = 1.080. The inverse theta of the NB model was 10.45. Considering that the NB distribution converges to the Poisson distribution when inverse theta approaches infinity and that the GOF and AIC for the Poisson and NB models were very similar we conclude that the Poisson model adequately describes the random distribution in this data.

Outbreak definition. Three observed WNV outbreaks were used to simulated the outbreaks in our model: French outbreaks in horses in 2000 [25] and 2004 [14] where 76 and 32 confirmed cases were reported respectively among 131 and 72 horses presenting nervous symptoms, and the Italian outbreak in 1998 [26] where 14 cases of WNV in horses were investigated by week of onset.

The weekly counts from these three outbreaks were fitted to the NB distribution. The resulting outbreak distribution was NB($\mu = 4.45$, $\theta = 0.94$). Based on this we predicted a median number of outbreak-related cases per week during an outbreak to be 3 with a 95% confidence interval of 0 to 18 cases.

Outbreak detection. Three scenarios were tested. The probability of an outbreak is not constant over the year, instead the relative probability of an outbreak occurring in spring (week 10 to 30), summer/autumn (week 31 to 46) and winter (week 47 to 9) is approximately 1:5:0.04. We chose to test one scenario per time period, i.e. the scenario A occurs in autumn, scenario B in winter and the scenario C in spring. For each scenario, the Poisson model was applied on the test set and one simulated peak/outbreak was inserted into the baseline (Figure 2). For each week the value of evidence was calculated using Eq5 where the probability of the observed number of cases during no outbreak $p(E|H_0)$ and during

The Value of Evidence from Syndromic Surveillance

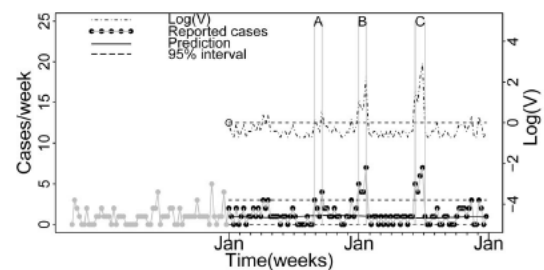


Figure 2. Application of NeurSy model on the test dataset. The vertical lines bounds peaks inserted during Year 1, week 36 to 39 (Scenario A), Year 2, week 1 to 4 (Scenario B) and Year 2, weeks 24 to 28 (Scenario C).

doi:10.1371/journal.pone.0111335.g002

outbreak $p(E|H_0)$ were calculated using the fitted model. Examples of the calculation of V during a non outbreak (scenario A) and during outbreaks (scenarios B and C) are shown in Figure 3.

Decision scenarios

The decision making in the outbreak scenarios for both examples is summarized in table 1.

The expected utility $u(C_0)$ for each scenario considered are given together with the action thresholds for posterior odds (O_{post}^*) and value of evidence (V^*) in favor of an outbreak. That is the situation for which the decision to act and not act have the same expected utility.

The expected utility of taking action in response to false alert ($u(C_{01})$) represents the costs for increased surveillance and preventive actions such as mosquito control for WNV. The utility of not taking action when there is an outbreak ($u(C_{10})$) represents the costs for control and economical and socio-economical consequences of an outbreak when the response to the outbreak was delayed. The losses may depend on season and in the example we have assumed that a WNV outbreak in summer or spring in the south of France results in extra costs due to its impact on tourism. Finally the utility of taking action when there is an outbreak ($u(C_{11})$) represents the costs for surveillance plus the economical and socio-economical impact in case of a timely response to the outbreak.

For NeurSy (Scenarios A to C), the prior odds in the table are based on the assumption that an outbreak of WNV is likely to occur every 3 years over an average of 5 weeks. The costs used are fictional but proportional to their expected relative contributions.

During the most at risk season regarding the probability of disease occurrence (Highest O_{pri}), the alarm threshold is low and 4 cases are sufficient to trigger an action (See Table 1. scenario A). For the season less at risk, the expected utilities are similar than during the most at risk season (O_{post}^* are equal), but no action is implemented even if 7 cases are reported because they are unlikely due to WNV (Low O_{pri}) (See Table 1. scenario B).

Sensitivity, specificity and receiver operating characteristics

The sensitivity and specificity of a surveillance system is defined by the chosen action threshold. The tradeoff between sensitivity and specificity of a model may be summarized in a receiver operating characteristics (ROC) curve [27]. The ROC curve corresponding to the case WNV case study is shown in figure 4A.

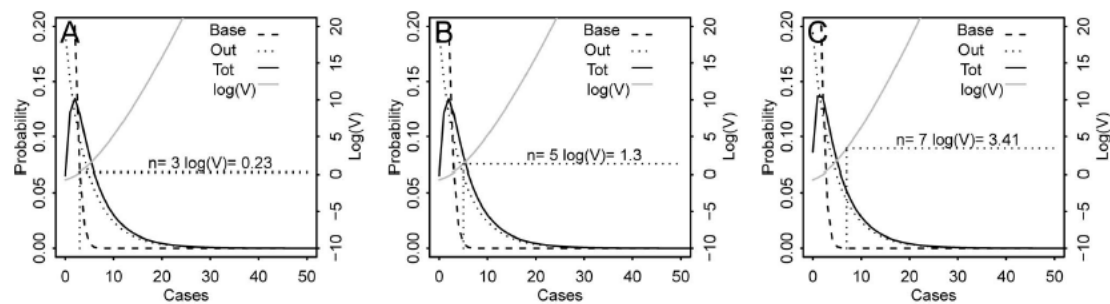


Figure 3. Value of evidence (V) and probability of observing n cases of neurological syndromes in a week during a non-outbreak (Base) situation and during a WNV outbreak (Out). Out is the distribution of outbreak related cases and Tot is the total number of observed cases per week during an outbreak. (A) Scenario A, year 1 week 36, $\lambda=1.08$, (B) Scenario B, year 2 week 1, $\lambda=1.08$, (C) Scenario C, year 2 week 27, $\lambda=0.81$.

doi:10.1371/journal.pone.0111335.g003

The values of SE and SP arising from scenarios A to C are indicated by letters. The PPV i.e. the probability that an alarm corresponds to a real outbreak [28] depends not only on SE and SP but also on the prior probability of an outbreak as indicated in figure 4B.

Case study 2– Respiratory syndromes and equine influenza (RespSy)

The same approach was successfully applied to the RespSy dataset. However, in this case the analysis indicated a significant degree of overdispersion in the weekly counts. Using the same regression model ($\text{counts} \sim \sin(2\pi t) + \cos(2\pi t) + \log(\text{histmean})$) the NB model had lower AIC (1141 vs 1284) and GOF closer to one (1.14 vs 2.54) compared to the Poisson model. The theta parameter for the NB distribution was 1.78, and resulting in a much wider confidence interval for the expected number of cases in a non-outbreak situation (Figure S2) compared to the Poisson model (Figure S3). When the NB and Poisson models are applied to the same test dataset (Figure S4, S5) the latter will report a value of evidence for the inserted peaks (D, E) that is several orders of magnitude higher than does the NB model. The Poisson model also reports peaks with $\text{Log}(V)$ close to 2 several times per year (Figure S5). An underlying assumption in the Poisson model is the

absence of overdispersion and, when this assumption does not hold, the Poisson model underestimates the probability of obtaining a large number of reported cases in the non-outbreak situation. Consequently it overestimates the value of evidence in favor of an outbreak. The overdispersion may be due to clustering in reporting. In the surveillance protocol veterinarians are encouraged not only to declare the diseased horse but also 1 to 3 additional horses (from the same stable), suspected to be in the incubation phase of influenza.

Discussion

In this work we have demonstrated how the value of evidence concept may be incorporated in a decision support system for syndromic surveillance and how the output may be used for risk assessment and informed decision making. According to the OIE - Terrestrial Animal Health Code [29] the decision to take action involves balancing costs for activities against economical and social consequences of a delayed response to an outbreak is the responsibility of the risk manager and should be separate from risk assessment.

Thus, although it is perfectly possible to build a system that outputs a best decision, the proposed approach is in concordance

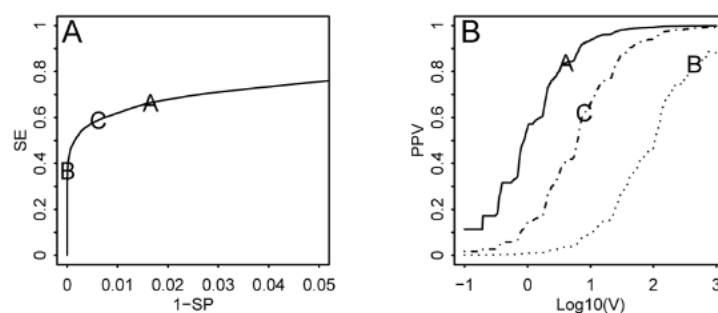


Figure 4. (A) ROC curve for outbreak detection of WNV based on neurological symptoms. Letters A–C indicate the decision threshold for $\text{Log}(V^*)$ in scenario A–C respectively (B). Positive Predictive Value (PPV) for different thresholds of $\text{Log}_{10}(V^*)$ given the prior probabilities of scenario A, B and C. The position of the letters indicate the action threshold for the respective scenario.

doi:10.1371/journal.pone.0111335.g004

with the risk analysis framework [29] by offering explicit separation of assumptions (P_{prior}), scientific evidence (V) and criteria for decisions and a transparency of how the evidence is evaluated. In forensics, the value of evidence is typically presented to the court as a qualitative statement in which fixed verbal expressions correspond to specified intervals for V [10,30]. This approach may be useful also when presenting epidemiological results. For example a value of $\text{Log}_{10}(V)$ in the range 1–2 may be expressed as “results provide moderate evidence to support that an outbreak is ongoing”. Alternatively intervals for V and/or O_{post} could be expressed using a color scale to produce maps representing the results from surveillance and risk of ongoing outbreaks of different diseases.

The model presented here is intended as a proof of concept and when setting up an operational syndromic surveillance system it will, as usual, be necessary to perform a careful evaluation of the baseline model to ensure that the regression model does not overfit to the baseline data. When designing the current model it was evident that high dimensional regression models were prone to find artefactual seasonal patterns that could severely bias the estimated probability of observing a number of counts in a particular week (results not shown). In the current implementation the model learns seasonal patterns and distribution of residuals (Inverse theta parameter of NB distribution) from manually curated data whereas the expected yearly average (*histmean*) is continuously updated from outbreak-filtered weekly data. Naturally the value of evidence concept may also be applied to a system where the baseline model is automatically retrained on new data. However, since the distribution parameter (theta) of the NB distribution would determine the cutoff in the filtering algorithm we argue that it is safer not to use the filtered data for estimation of the same parameter without prior inspection of the data. The same conclusion holds for seasonal patterns.

The overdispersion in the RespSy dataset is largely due to veterinarians sampling several horses in a stable upon suspicion. Thus, in this special case it might be possible to handle the overdispersion by pre-processing the data to remove redundant cases, provided that the same pre-processing is applied to new data on weekly basis. However, when the mechanism behind overdispersion in baseline counts is not so transparent that automatic filtering out redundant cases is possible the NB model will support a correct interpretation of the value of the peak in the count data.

As indicated in Figure 4 the tradeoff between SE and SP differs between seasons. This is natural since in case the (prior) probability of an outbreak differs between seasons the average sensitivity SE_{avr} and specificity SP_{avr} will be given by:

$$SE_{avr} = \frac{\sum_{i=1}^n (d_i \times SE_i \times P_i)}{\sum_{i=1}^n (d_i)} \quad (\text{Eq.16})$$

$$SP_{avr} = \frac{\sum_{i=1}^n (d_i \times SP_i \times (1 - P_i))}{\sum_{i=1}^n (d_i)} \quad (\text{Eq.17})$$

where:

SE_i = sensitivity in season i

SP_i = specificity in season i

P_i = (prior) probability of outbreak in season i

d_i = (relative) duration of season i

Thus, by incorporating prior knowledge about the seasonality of the diseases of interest it is possible to achieve a high average sensitivity without sacrificing the PPV and SP. Another important attribute of outbreak detection is timeliness. Whereas there is no

general measure of timeliness [28] the number of cases are often small in the first week(s) of an outbreak, increasing the sensitivity (i.e. lowering the threshold for V and thus n) in the high risk season will result in improved timeliness as well as average sensitivity.

In this work we have introduced the framework using models that evaluate evidence from each week independently. Although this simple approach is suitable for presenting the framework and a reasonable choice for an early warning system, the evaluation of evidence from one week at a time is not a fundamental limitation of the approach. A model accounting for accumulation of evidence over several weeks may, for example, be constructed by considering, for each week in the interval $[0...j]$ the conditional probability

$$P(E|H_{i,t})$$

Where

t is the week of interest

$H_{i,t}$ is the hypothesis that an outbreak started i weeks before t

E_{t-n} is the number of reported cases in week $[t-i... t]$

The probability of observing n outbreak-related cases will not be uniform throughout the outbreak but depend on whether the outbreak is in its first, second or third week *etc.* When accounting for evidence from several weeks the value of evidence in favor of the hypothesis H_1 “An outbreak is going on” against H_0 “An outbreak is not going on” will be dependent on the prior probability of an outbreak starting in any of the preceding weeks. This is due to the fact that H_1 is composed of several sub-hypotheses:

$H_{1,i=0}$: An outbreak started in week t

$H_{1,i=1}$: An outbreak started in week $t-1$

..

$H_{1,i=j}$: An outbreak started in week $t-j$

Consequently $p(H_1)$ depends on the relative probability of these sub-hypotheses. The value of evidence in favor of an outbreak going on in week of interest (V) can be calculated as the Bayes factor (B):

$$V = B = \frac{O_{post}}{O_{pri}} \quad (\text{Eq.18})$$

where

O_{post} is the posterior odds of an outbreak going on in week of interest

O_{pri} is the prior odds of an outbreak going on in week of interest

Although in the more complex models the calculation of the value of evidence would depend on the prior probability of outbreak, the framework is still applicable for communicating the evidence to decision makers. Essentially any Markov Chain model could be applied in the evaluation of evidence framework and the choice of complexity is a tradeoff between on the one hand realism and on the other hand simplicity and transparency. However, we anticipate that in most situations there will not be sufficient data to support very complex models.

Supporting Information

Figure S1 Fitted baseline and one sided 95% confidence interval for weekly counts for case NeurSy Years 2006–2010. Poisson regression using model: counts $\sim \sin(2\pi t) + \cos(2\pi t) + \log(\text{histmean})$. (TIF)

Figure S2 Fitted baseline and one sided 95% confidence interval for weekly counts for case RespSy Years 2006–2010. NB regression using model: counts $\sim \sin(2\pi t) + \cos(2\pi t) + \log(\text{histmean})$.
(TIF)

Figure S3 Fitted baseline and one sided 95% confidence interval for weekly counts for case RespSy Years 2006–2010. Poisson regression using model: counts $\sim \sin(2\pi t) + \cos(2\pi t) + \log(\text{histmean})$.
(TIF)

Figure S4 Application of RespSy NB-model on the fictive test dataset. The vertical lines bounds peaks inserted during Year 1, week 36 to 39 (D), Year 2, week 24 to 28 (E). The gray points indicate historical data used to calculate the historical average (histmean).
(TIF)

Figure S5 Application of RespSy Poisson-model on the fictive test dataset. The vertical lines bounds peaks inserted during Year 1, week 36 to 39 (D), Year 2, week 24 to 28 (E). The gray points indicate historical data used to calculate the historical average (histmean).
(TIF)

Table S1 NeurSy baseline 2006–2012.
(CSV)

Table S2 RespSy baseline 2006–2012.
(CSV)

Table S3 NeurSy outbreak distribution.
(CSV)

Table S4 RespSy outbreak distribution.
(CSV)

Script S1 Script used to analyze time series data.
(R)

Script S2 Script used to illustrate the calculation of V in figures 1 and 3.
(R)

Script S3 Script used to prepare ROC and PPV plots.
(R)

Script S4 Script used to remove aberrations from baseline when preparing fictive test data.
(R)

Acknowledgments

Christel Marcillaud-Pitel for collecting and providing data from the RESPE (Réseau d'Epidémiologie et Pathologie Equine) and Fernanda Dórea for help with dynamic regression models.

Author Contributions

Conceived and designed the experiments: MGA. Analyzed the data: MGA CF. Contributed reagents/materials/analysis tools: AL LL. Wrote the paper: MGA CF FV AL. Organized baseline data collection: AL LL. Critical revision of manuscript: LL.

References

- Hillier KM, Stoneking L, Min A, Rhodes SM (2013) Syndromic Surveillance for Influenza in the Emergency Department-A Systematic Review. *PLoS One* 8: e73832.
- Gesualdo F, Silo G, Agricola E, Gonfiantini MV, Pandolfi E, et al. (2013) Influenza-Like Illness Surveillance on Twitter through Automated Learning of Naïve Language. *PLoS One* 8: e82489.
- Gibbs EP (2014) The evolution of One Health: a decade of progress and challenges for the future. *Vet Rec* 174: 85–91.
- Dórea FC, Sanchez J, Revie CW (2011) Veterinary syndromic surveillance: Current initiatives and potential for development. *Prev Vet Med* 101: 1–17.
- Katz R, May L, Baker J, Test E (2011) Redefining syndromic surveillance. *Journal Epidemiol Global H* 1: 21–31.
- Shmueli G, Burkom H (2010) Statistical Challenges Facing Early Outbreak Detection in Biosurveillance. *Technometrics* 52: 39–51.
- Buchler JW, Berkelman RL, Hartley DM, Peters CJ (2003) Syndromic Surveillance and Bioterrorism-related Epidemics. *Emerg Infect Dis* 9: 1197–1204.
- Brownson RC, Fielding JE, Maylahn CM (2009) Evidence-Based Public Health: A Fundamental Concept for Public Health Practice. *Annu Rev of Publ Health* 30: 175–201.
- Taroni F, Aitken C, Garbolino P, Biedermann A (2006) Bayesian Networks and Probabilistic Inference in Forensic Science. Chichester: John Wiley & Sons Ltd.
- Aitken C, Roberts P, Jackson G (2010) Fundamentals of Probability and Statistical Evidence in Criminal Proceedings. Guidance for Judges, Lawyers, Forensic Scientists and Expert Witnesses. London: The Royal Statistical Society.
- Jarman KH, Kreuzer-Martin HW, Wunschel DS, Valentine NB, Clift JB, et al. (2008) Bayesian-integrated microbial forensics. *Appl and Environ Microb* 74: 3573–3582.
- Keats A, FS L, Yee E (2006) Source Determination in Built-Up Environments Through Bayesian Inference With Validation Using the MUST Array and Joint Urban 2003 Tracer Experiments. *Proc 14th Annual Conference of the Computational Fluid Dynamics Society of Canada*, July 16–18, Kingston, Canada.
- Andersson G, Aspán A, Hultén C, Ågren E, Barker GC (2013) Application of forensic evaluation of evidence to the tracing of Salmonella. *Symposium Salmonella and Salmonellosis 13S*. Saint-Malo FRANCE.
- Leblond A, Hendriks P, Sabatier P (2007) West Nile Virus Outbreak Detection Using Syndromic Monitoring in Horses. *Vector-Borne Zoonot* 7: 403–410.
- Legrand IJ, Pitel PH, Marcillaud-Pitel CJ, Cullinane AA, Courouche AM, et al. (2013) Surveillance of equine influenza viruses through the RESPE network in France from November 2005 to October 2010. *Equine vet j* 45: 776–783.
- Martella V, Elia G, Decaro N, Di Trani L, Lorusso E, et al. (2007) An outbreak of equine influenza virus in vaccinated horses in Italy is due to an H3N8 strain closely related to recent North American representatives of the Florida sub-lineage. *Vet Microbiol* 121: 56–63.
- Newton JR, Daly JM, Spencer L, Mumford JA (2006) Description of the outbreak of equine influenza (H3N8) in the United Kingdom in 2003, during which recently vaccinated horses in Newmarket developed respiratory disease. *Vet Rec* 158: 185–192.
- Yamanaka T, Niwa H, Tsujimura K, Kondo T, Matsumura T (2008) Epidemic of equine influenza among vaccinated racehorses in Japan in 2007. *J Vet Med Sci* 70: 623–625.
- Schmidt AM, Pereira JBM (2011) Modelling Time Series of Counts in Epidemiology. *Int Stat Rev* 79: 48–69.
- Cameron AR (2012) The consequences of risk-based surveillance: Developing output-based standards for surveillance to demonstrate freedom from disease. *Prev Vet Med* 105: 280–286.
- Gittelsohn SN (2013) Evolving from inferences to decisions in the interpretation of scientific evidence. *Institute de Police Scientifique*. Lausanne: Université de Lausanne.
- R Core Team (2013) R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.
- Venables WN, Ripley BD (2002) Modern Applied Statistics with S. New York: Springer.
- Burkom HS, Elbert Y, Feldman A, Lin J (2004) Role of data aggregation in biosurveillance detection strategies with applications from ESSENCE. *MMWR Morbidity and mortality weekly report* 53 Suppl: 67–73.
- Murgue B, Murri S, Zientara S, Durand B, Durand JP, et al. (2001) West Nile outbreak in horses in southern France, 2000: the return after 35 years. *Emerg Infect Dis* 7: 692–696.
- Autorino GL, Battisti A, Deubel V, Ferrari G, Forletta R, et al. (2002) West Nile virus Epidemic in Horses, Tuscany Region, Italy. *Emerg Infect Dis* 8: 1372–1378.
- Hastie T, Tibshirani R, Friedman J (2001) The elements of statistical learning, Data mining, inference and Prediction. New York: Springer Science + Business Media Inc.
- Sosin DM (2003) Draft framework for evaluating syndromic surveillance systems. *J Urban Health* 80: i8–13.
- OIE (2011) OIE - Terrestrial Animal Health Code. World Organization for Animal Health.
- Nordgaard A, Ansell R, Drotz W, Jaeger L (2012) Scale of conclusions for the value of evidence. *Law, probability & risk* 11: 1–24.

D. DISCUSSION AND CONCLUSION

1. Discussion

This chapter showed that syndromic surveillance based on data collected by RESPE is able to detect signals suggestive of the presence of an outbreak of WNV or equine influenza in French horses (through the monitoring of nervous and respiratory symptoms, respectively). Although the system appears to be effective, syndromic surveillance is still rarely implemented in horses, probably due at least in part to the lack of data available for this population (e.g., few centralized databases on horse health, mainly due to the diversity of activities in the equine industry). In Europe, we have only been able to identify three syndromic surveillance initiatives for horses, based on a recent review by Dupuy *et al.* (Dupuy *et al.* 2013a) and an additional literature search. One such initiative is present in the UK, associated with the Equine Quarterly Surveillance Reports (DEFRA/AHS/BEVA 2015), another is in the Netherlands (Rockx *et al.* 2006), and the third is in Switzerland, linked to the Equinella network (<https://www.equinella.ch/>). At the time of writing, none of these systems (RESPE included) uses statistical tools to detect the signal of an outbreak prior to its formal diagnosis. Instead, the alarm threshold is typically set at an arbitrary level using raw data. However, our results indicate that data preprocessing and analysis improve detection performance compared to such subjective methods.

In syndromic surveillance, Bayesian approaches have previously been mainly used for spatiotemporal outbreak detection and/or to assess unknown posterior probabilities by using hierarchical Bayesian models that involve inferences (Chan *et al.* 2010, Banks *et al.* 2012, Neill *et al.* 2006, Schmidt and Pereira 2011, Zou *et al.* 2010). The simple Bayesian approach developed in this chapter is less robust, from a mathematical point of view, than full inference-based Bayesian models, which take into account the uncertainties of parameter estimation. However, our approach is also much easier to implement and to understand. This is an advantage compared to more complex approaches, which are often too complicated for decision makers who lack experience with these methods to understand (Banks *et al.* 2012). The simple Bayesian approach used here could thus be a good compromise between rigor and ease-of-understanding in presenting results.

In this chapter, two types of detection algorithms were tested: classical methods based on regression models and an algorithm based on the empirical Bayesian approach. Classical approaches generated good detection performance and were simple to apply, a great advantage in veterinary medicine where syndromic surveillance systems are still difficult to implement (Shephard 2006). However, compared to an empirical Bayesian approach, classical approaches have three main shortcomings: they do not provide a quantitative output, they are not able to easily take into account other epidemiological information available on a disease, and they are very unspecific. This last point, however, can be an advantage in detecting an unknown disease, as no *a priori* hypotheses are required. The empirical Bayesian approach is more disease-specific, but the drawback is that knowledge of the disease must be available. However, this

is not a fundamental limitation to the approach, as a vague probability distribution can also be used. The choice of an approach should be made according to the objective of the surveillance and the data available. In addition, it would be useful to perform a formal and quantitative comparison of the detection performances of both approaches before reaching a conclusion about their advantages and drawbacks.

The representativeness of the RESPE data was not evaluated here, but was recently investigated by that organization (Daix 2014). The study identified 430 sentinel veterinarians who truly participated in the reporting system. They are not equally distributed all over France, but rather present in 92 (of 96) French regions and concentrated in areas with high horse densities. The questionnaires filled out by 63% of the active sentinel veterinarians revealed that 26% do not declare all suspect cases to RESPE. The reasons cited were: no consent from owner, omission, lack of time, and definitive diagnosis obtained without laboratory analysis. Most of the sentinel veterinarians reported seeing only a few suspect cases, which explained the low number of declarations reported per veterinarian. These elements suggest that the representativeness of the RESPE data is probably acceptable. However, it is unknown how these factors specifically affect the reporting of nervous and respiratory symptoms. Indeed, some veterinarians seem to declare only certain symptoms and not others. Therefore, it would be important to perform a detailed assessment of the RESPE system in order to ensure optimal outbreak detection performance.

2. Conclusion

This chapter describes a useful tool for determining if there is an ongoing VBD outbreak in French horses. In particular, the Bayesian approach enabled us to merge syndromic surveillance with knowledge of risk factors, which can be especially useful for VBDs as they are closely related to season and environment.

Despite these advantages, an outbreak may also occur without generating a detectable signal in the syndromic surveillance system. This is of particular concern given the unknown representativeness of the data used. In addition, even if there is a significant signal, the lack of specificity in the system means that it could be a false alarm. It will thus be useful going forward to consider other approaches for estimating the probability of an outbreak in order to ensure the early detection of a newly introduced disease.

CHAPTER IV: MULTIPLE INDICATORS OF RISK

Risk assessments and syndromic surveillance provide an estimate of the probability of an outbreak, but they do not prove the presence of the disease. The accumulation of evidence from a combination of these approaches can be a way to improve confidence in the prediction that a newly introduced epidemic is in progress. More generally, combining all available information in a multivariate algorithm should give better results for outbreak detection than univariate methods do.

In this chapter, we explore the combination of multiple risk indicators in order to improve assessment of the probability of occurrence of a newly introduced VBD in horses. We first direct our attention to WNV, which infects a wide range of species and induces different types of symptoms. Therefore, combining information from different species or from different syndromic groups could help in outbreak detection. We thus worked first on a multivariate syndromic surveillance system in order to improve the detection of WNV outbreaks (Chapter IV.A). However, as already highlighted, VBDs are closely linked to season and environment. In addition, the probability that an outbreak occurs is also linked to the probability that the disease enters an area. Combining syndromic surveillance with assessments of the probabilities of virus entry and establishment may thus also improve confidence in the prediction that a new VBD has been introduced. This approach has the potential to be much more useful than syndromic surveillance alone, as the latter may fail to detect an outbreak of a disease that causes few symptoms. To test this approach, we worked to combine syndromic surveillance data with quantitative risk analysis and we applied this approach to the detection of equine encephalosis (EE) outbreaks. We chose EE because it has few symptoms and should be more difficult to detect using a single approach than AHS and WNV, both of which manifest themselves in a more specific clinical picture (Chapter IV.B). The advantages and drawbacks of combining these risk indicators to detect VBDs in horses are also discussed (Chap. IV.C).

A. MULTISTREAM SYNDROMIC SURVEILLANCE

1. Introduction

One limit of univariate syndromic surveillance is that no single data source captures all the individuals involved in the outbreak, and that diseases may cause a wide variety of symptoms in different individuals (Kulldorff et al. 2007). In addition, the data collected are often vague, and univariate syndromic surveillance systems can suffer from this lack of specificity. However, by simultaneously assessing information from different data sources related to different populations and/or symptoms, one can improve outbreak detection and, in particular, the specificity of the detection.

Multivariate syndromic surveillance can be purely temporal (see, for example, (Schiöler and Frisé 2011), (Fan et al. 2014), and (Lau et al. 2008)), purely spatial, or spatiotemporal (see, for example, (Kulldorff et al. 2007) and (Greene et al. 2012)). Different methods exist for the aggregation of data sources, but the two main approaches are the reduction method and the parallel method (Frisé et al. 2010; Sonesson and Frisé 2005). The reduction method considers several variables that are reduced to a single statistic, using for example the sum for each time period, p-value aggregation (Roure et al. 2007), or multivariate control charts (MacGregor and Kourti 1995; Stoto et al. 2006)). The parallel method uses multiple univariate systems which are then assessed in parallel. An alarm is triggered if any of the univariate systems gives an alarm (e.g., (Fan et al. 2014, Schiöler and Frisé 2011)). The reduction approach gives a better detection performance when all changes occur simultaneously in the different processes under consideration. When the changes occur separately, the parallel approach yields better results (Frisé et al. 2010).

In this section, we consider a multivariate syndromic surveillance system applied to the detection of West Nile virus outbreaks. WNV typically affects different hosts, which makes it a particularly interesting case in determining if multivariate surveillance can improve outbreak detection. To combine the different variables, we used a reduction method based on a Bayesian approach.

This work resulted in a scientific paper submitted to *Vector-borne and zoonotic diseases* (see below for the main text).

2. Paper 3

TITLE: Evaluation of a multivariate syndromic surveillance system for West Nile virus.

AUTHORSHIP: C. Faverjon¹, M.G. Andersson², A. Decors³, J. Tapprest⁴, P. Triz⁵, A. Sandoz^{6,7}, O. Kutasi⁸, C. Sala⁹, A. Leblond¹⁰

- (1) INRA UR0346 Animal Epidemiology, VetagroSup, F-69280 Marcy l'Etoile, FRANCE;
- (2) Department of Chemistry, Environment and Feed Hygiene. The National Veterinary Institute, SE- 751 89 Uppsala, SWEDEN;
- (3) Office National de la Chasse et de la Faune Sauvage, Direction des études et de la recherche, 5 rue de Saint Thibaud, 78610 Auffargis, France.
- (4) ANSES Dozulé Laboratory for Equine Diseases, FRANCE.
- (5) Clinique vétérinaire, 19 Rue de Créhange, 57380 Faulquemont, collège syndrome nerveux du RESPE et Commission maladies infectieuses de l'AVEF, 6 avenue du Maréchal Montgomery 14000 Caen, France.
- (6) Centre de recherche pour la conservation des zones humides méditerranéennes, Fondation Tour du Valat, Arles F-13200, FRANCE.
- (7) UFR Sciences, Aix-Marseille Université, Marseille F-13003, France.
- (8) Hungarian Academy of Sciences - Szent Istvan University (MTA-SZIE) Large Animal Clinical Research Group, 2225 Ullo, Dóra major, HUNGARY.
- (9) ANSES-Lyon, Epidemiology Unit, 69364 Lyon, France
- (10) INRA UR0346 Animal Epidemiology et Département Hippique, VetAgroSup, F-69280, Marcy l'Etoile, FRANCE; and Réseau d'Epidémiologie et de Surveillance en Pathologie Equine (RESPE), Caen, FRANCE.

CORRESPONDENCE ADDRESS

FAVERJON Céline, celine.faverjon@vetagro-sup.fr, INRA UR0346 Animal Epidemiology, VetagroSup, F-69280 Marcy l'Etoile, FRANCE

ABSTRACT

Background: Various methods are currently used for the early detection of West Nile virus (WNV) but their output is either not quantitative or does not take into account all available information. Our study aimed to test a multivariate syndromic surveillance system in order to improve early detection of WNV.

Method: Weekly time series data on nervous syndromes in horses and mortality in both horses and wild birds were used. Baselines were fitted to the three time series and used to simulate 100 years of surveillance data. WNV outbreaks were simulated and inserted into the baselines based on historical data and expert opinion. Univariate and multivariate syndromic surveillance systems were tested in order to gauge how well they detected the outbreaks; detection was based on an empirical Bayesian approach. The systems' performances were compared using measures of sensitivity, specificity, and area-under-ROC-curve (AUC).

Result: When data sources were considered separately (i.e. univariate systems), the best detection performance was obtained using the dataset of nervous symptoms in horses compared to those of bird and horse mortality (AUCs respectively equal to 0.80, 0.75, and 0.50). A multivariate outbreak detection system that used nervous symptoms in horses and bird mortality generated the best performance (AUC = 0.87).

Conclusion: The proposed approach is suitable for performing multivariate syndromic surveillance of WNV outbreaks. This is particularly relevant given that a multivariate surveillance system performed better than a univariate approach. Such a surveillance system could be especially useful in serving as an alert for the possibility of human viral infections. This approach can be also used for other diseases for which multiple sources of evidence are available.

KEY WORDS: West Nile, syndromic surveillance, Bayes, horses, multivariate detection

ACKNOWLEDGMENTS

Data on mortality in horses were provided by the French Ministry of Agriculture. Data on mortality in wild birds were provided by the event-based surveillance system "SAGIR", the French surveillance network for wild birds and mammals. Data on nervous symptoms in horses were provided by 'RESPE' (Réseau d'Epidémio-Surveillance en Pathologie Equine). The authors wish to thank all persons involved in these networks (hunters, technicians of FDC and environmental inspectors of ONCFS, local veterinary laboratories, veterinarians, and fallen stock companies) for collecting these data and all these institutes for providing them.

This research was carried out as part of the VICE project funded under the FP7 EMIDA ERA-NET initiative. National funding within ERA-Net was provided by the Dutch Ministry of Economic Affairs (Project nr. BO-20-009-009) and by the Swedish research council Formas. This work was also partially financially supported by IFCE (Institut Français du Cheval et de l'Equitation) and Vetagro Sup (National Veterinary School of Lyon).

INTRODUCTION

West Nile virus (WNV) is a mosquito-borne arbovirus mainly transmitted by mosquitoes from the genus *Culex* (family Culicidae). Its main hosts are birds but the virus also affects various non-avian species including horses and humans, with dramatic consequences for public health and for the equine industry, i.e. potentially fatal encephalitis in humans and horses (Campbell et al. 2002; Castillo-Olivares and Wood 2004). In Europe, WNV emerged in the 1960s and several outbreaks have been documented since that time (Calistri et al. 2010). Even if the virus is now considered endemic in a large part of Europe, the number of reported outbreaks is presently increasing in southern and eastern Europe, particularly in Italy, Greece, and Bulgaria (Di Sabatino et al. 2014). This increasing number of outbreaks, combined with the recent introduction and spread in Europe of WNV lineage 2, which induces severe symptoms in humans, horses, and birds (Bakonyi et al. 2006; Calzolari et al. 2013; Hernández-Triana et al. 2014), has resulted in growing concern about WNV in Europe. In addition, the implementation of prevention plans for WNV outbreaks is difficult (Zeller 2010) because the environmental factors and meteorological interactions underlying the increase in WNV circulating in mammals are still poorly understood. To improve early detection of WNV outbreaks, then, the major challenge is to develop more integrated and quantitative approaches (Beck et al. 2013; Bellini et al. 2014b).

Syndromic surveillance is currently a popular approach for the early detection of health-related phenomena (Dórea et al. 2011) and has already been implemented for WNV. In Europe, the surveillance of nervous syndromes in horses has been shown to detect early indicators of WNV outbreaks (Leblond et al. 2007; Saegerman et al. 2014) and is one of the most cost-effective surveillance systems in the European context (Chevalier et al. 2011). In the USA, instead, increased mortality in wild birds is one of the most timely indicators of virus activity (Brown 2012). Mortality in wild birds had rarely been reported in Europe until the recent explosive spread of lineage 2 in 2008-2009 in Hungary and Austria, which suggests that this parameter could be also incorporated into future monitoring systems in Europe (Bakonyi et al. 2013). This is consistent with recent experimental infections of European wild birds with various WNV strains, which generated an average mortality rate of 43% (Del Amo et al. 2014a; Del Amo et al. 2014b; Dridi et al. 2013; Sotelo et al. 2011; Ziegler et al. 2013). Apart from mortality in wild birds and nervous symptoms in horses, WNV is also associated with mortality in horses, which could constitute another signal of a WNV outbreak. Combining all available information in a multivariate algorithm should give better results for outbreak detection than univariate methods alone. However, at the time of writing, multivariate syndromic surveillance has never been implemented for the detection of WNV outbreaks.

The aim of our study was to evaluate the performance of a multivariate syndromic surveillance system in detecting WNV using three datasets: nervous syndromes in horses and mortality in horses and wild birds. We focused on the French Mediterranean coast, which is a particularly high-risk area for WNV outbreaks. Indeed, in France, WNV has only ever been identified in this area, which is home to mammalian and avian hosts, bridging vectors, and large protected wetlands with numerous migratory birds.

MATERIALS AND METHODS

1. Data sources

1.1. Nervous syndromes in horses

Data on nervous syndromes in horses are collected through the passive surveillance system “RESPE”. This French network for the surveillance of equine diseases (<http://www.respe.net/>) collects standardized declarations from veterinary practitioners registered as sentinels. All the samples sent for laboratory diagnosis are systematically tested for WNV and equine herpes virus, and results are registered in the RESPE database. To obtain an outbreak-free baseline dataset, we used data from 2006 to 2013 that included only the 44 declarations without positive laboratory test results from the region of the French Mediterranean coast. The time series of nervous syndromes in horses is designated *NervSy* in subsequent sections.

1.2. Mortality in horses

Data on mortality in horses have been centralized since 2010 in the “EDI-SPAN” database, managed by all the French fallen stock companies and the French Ministry of Agriculture (Perrin et al. 2012). As WNV does not produce perinatal mortality, we only considered the 8 742 dead adult horses collected around the French Mediterranean coast between 2010 and 2014. The time series of mortality in adult horses is designated *DeadHorse* in subsequent sections.

1.3. Mortality in wild birds

Data on mortality in wild birds are collected through the event-based surveillance system “SAGIR”, the national French surveillance network of diseases in wild birds and mammals, which collects declarations from field workers (e.g., hunters, technicians from departmental hunting federations, and environmental inspectors from the French National Hunting and Wildlife Agency (ONCFS)). Surveillance relies on diagnosis at a local veterinary laboratory (Decors et al. 2014). Between 2007 and 2013, 292 dead wild birds were collected and necropsied around the French Mediterranean coast. The time series of the number of necropsied wild birds is designated *DeadBird* in subsequent sections.

2. Data modeling and simulation

2.1. Baselines

All time series were aggregated weekly. Using visual examination, abnormal peaks were observed only in *DeadBird*. These extreme values were removed based on a method adapted from Tsui *et al.* (Tsui et al. 2001): the entire dataset was first fitted to a Poisson distribution and then values above the 95% confidence interval were deleted and replaced with the average value of the four previous weeks.

To calibrate the models, we used *NervSy* data from 2006 to 2010, *DeadHorse* data from 2011 to 2013, and *DeadBird* data from 2007 to 2011. Instead, to validate the quality of predictions, we used *NervSy* data from 2011 to 2013, *DeadHorse* data from 2014, and *DeadBird* data from 2012 to 2013. To define the background noise of the time series without outbreaks, we fitted alternative regression models based on

Poisson and negative binomial (\mathcal{NB}) distributions. Models were implemented in R x64 version 3.0.2. Dynamic regression was performed with the functions *glm* (package {stats}) and *glm.nb* (package {MASS}). The expected number of counts at time t was estimated with the *predict* functions of the respective packages.

Models were evaluated using the Akaike information criterion (AIC) (Bozdogan 1987), and the adjusted deviance (deviance/degree of freedom) was used as a measure of goodness-of-fit (GOF). The agreement between predicted and observed values was assessed according to the root-mean-squared error (Chai and Draxler 2014). The criterion was assessed within the calibration period ($RMSE_c$) and within the validation period ($RMSE_v$). In either case, the lower the value, the better the predictive performance of the model.

For each time series, the best regression model was used to predict the expected value of each week of the next simulated year. Distribution of cases for each week was defined as a Poisson distribution with lambda equals to the predicted value for the same week. Weekly samples from 100 fictive years were generated by random sampling from the previous distributions as proposed by Dórea *et al.* (Dórea *et al.* 2013a).

2.2. WNV outbreaks

Data on real WNV outbreaks are scarce, so we thus used simulated outbreaks to evaluate our detection system. For each syndrome, the distribution of the number of cases during an outbreak was estimated with the *fitdist* function of the package {fitdistrplus}. Time series for each syndrome during 100 fictive outbreaks were simulated by randomly sampling the corresponding distribution. One simulated outbreak was inserted in each simulated baseline. The outbreaks related to nervous cases in horses were randomly inserted, followed by the corresponding outbreaks related to wild bird mortality, such that the time lag between the first dead bird and the first nervous case in horses due to WNV was 0, 1, or 2 weeks (Kulasekera *et al.* 2001). The corresponding horse mortality outbreaks were inserted such that half of the affected horses died the week of onset of clinical signs and half died the week after (Bunning *et al.* 2002; Cantile *et al.* 2000; Trock *et al.* 2001; Ward *et al.* 2006).

The weekly counts of cases of five real European WNV outbreaks (Anonymous 2007; Autorino *et al.* 2002; Kutasi *et al.* 2011; Leblond *et al.* 2007; Murgue *et al.* 2001) were fitted to the \mathcal{NB} distribution and the resulting distribution of the additional number of nervous cases due to WNV during an outbreak was $\mathcal{NB}(\mu=3.12, \theta=1.150)$. The mortality among horses clinically affected by WNV was fitted to a normal distribution (mean=0.384, standard deviation=0.128) based on (Autorino *et al.* 2002; Leblond *et al.* 2007; Murgue *et al.* 2001; Ward *et al.* 2006). The *NervSy* dataset did not provide the real number of clinically affected horses, so we assumed that only 50% of horses with nervous symptoms were declared to RESPE. To estimate the real number of clinically affected horses, we simulated RESPE declarations of nervous symptoms associated with 100 WNV outbreaks and doubled the counts of horses obtained. The related weekly count of dead adult horses was then deduced and fitted to the \mathcal{NB} distribution $\mathcal{NB}(\mu=3,$

theta=2.005). The distribution of the weekly number of dead birds was estimated by expert opinion to be $\mathcal{NB}(\text{mean}=2.23, \text{theta}=3.34)$.

3. Outbreak detection

3.1. Bayesian framework

Bayesian hypothesis testing is based on two mutually exclusive hypotheses which can be expressed in the syndromic surveillance context as H_1 , “there is an ongoing outbreak of WNV (or another disease with similar symptoms)”, and H_0 , “there is no ongoing outbreak” (Andersson et al. 2014). The relative probability of the two hypotheses can be expressed as a ratio (O_{pri}) which represents our *a priori* belief about the disease status:

$$\text{Eq.1} \quad O_{pri} = \frac{P(H_1)}{P(H_0)}$$

When evidence in favor (or not) of each hypothesis is observed, we can build the *a posteriori* belief about the disease's status (O_{post}):

$$\text{Eq.2} \quad O_{post} = \frac{P(H_1 | E_x)}{P(H_0 | E_x)}$$

where $P(H_1 | E_x)$ is the probability of H_1 given the evidence E observed in time series x and $P(H_0 | E_x)$ is the probability of H_0 given the evidence E observed in time series x .

Using this general framework with the application of Bayes' theorem, O_{post} can be calculated as:

$$\text{Eq.3} \quad O_{post} = V_x \times O_{pri} = \frac{P(E_x | H_1)}{P(E_x | H_0)} \times \frac{P(H_1)}{P(H_0)}$$

where V_x is the value of evidence, $P(E_x | H_1)$ is the probability of observing the number of reported cases of syndrome x in a particular week given that H_1 is true, and $P(E_x | H_0)$ is the probability of observing the number of reported cases of syndrome x in a particular week given that H_0 is true.

In order to estimate $P(E_x | H_1)$ and $P(E_x | H_0)$, information on the probability distribution for the number of reported cases in non-outbreak and outbreak situations is used. The probability of E_x (observation of n cases in time series x) during an outbreak is calculated as:

$$\text{Eq.4} \quad P(E_x | H_1) = \sum_{i=0}^n P_{base}(i) \times P_{out}(n-i)$$

where $P_{base}(i)$ is the probability of drawing i cases from the baseline distribution in time series x and $P_{out}(i)$ is the probability of drawing i cases from the outbreak distribution in time series x based on the shape of the outbreak, as previously simulated.

3.2. Combining time series

When the three time series were combined, V_{tot} incorporated evidence from *NervSy*, *DeadHorse*, and *DeadBird*, respectively denoted as E_{NervSy} , $E_{DeadHorse}$, and $E_{DeadBird}$. Assuming that the three sources of evidence were independent, V_{tot} was calculated as:

$$Eq.5 \quad V_{tot} = \frac{P(E_{NervSy} E_{DeadHorse} E_{DeadBird} | H_1)}{P(E_{NervSy} E_{DeadHorse} E_{DeadBird} | H_0)} = V_{NervSy} \times V_{DeadHorse} \times V_{DeadBird}$$

and O_{post_tot} was calculated as:

$$Eq.6 \quad O_{post_tot} = \frac{P(H_1 | E_{NervSy} E_{DeadHorse} E_{DeadBird})}{P(H_0 | E_{NervSy} E_{DeadHorse} E_{DeadBird})} = V_{tot} \times \frac{P(H_1)}{P(H_0)}$$

4. Performance assessment

Sensitivity (Se) and specificity (Sp) were calculated as:

$$Eq.7 \quad Se = TP / (TP + FN)$$

$$Eq.8 \quad Sp = TN / (TN + FP)$$

where TP is the number of true positive alarms, TN the number of true negative alarms, FP the number of false positive alarms, and FN the number of false negative alarms.

The receiver operating characteristic (ROC) curve was generated in R by testing various alarm thresholds, and the areas under the curves (AUC) were calculated with the *auc* function of the package {flux}. A larger AUC represented a better detection performance.

RESULTS

1. Modeling time series and simulating data

For all time series the best fits were obtained for \mathcal{NB} distributions. The resulting models' parameters are summarized in table 1 and corresponding baselines and predictions are shown in figure 1. The probabilities of observing n cases and the resulting value of V ($p(E|H_1)/p(E|H_0)$) during a non-outbreak ($p(E|H_0)$) and an outbreak ($p(E|H_1)$) situation for each time series are summarized in figure 2.

2. Outbreak detection

We estimated the respective performance of each univariate system (*NervSy*, *DeadHorse*, and *DeadBird*) in detecting WNV outbreaks without considering any *a priori* values for disease status ($O_{pri}=1$). Examples of simulated baselines with inserted outbreaks and associated variations in $\log_{10}(V)$ are presented in Appendix I.

The best results for univariate outbreak detection were obtained for *NervSy*, which outperformed analyses using *DeadHorse* and *DeadBird* (figure 3 and table 2). *DeadBird* models yielded intermediary

detection performances whereas models using *DeadHorse* were not able to discriminate between outbreak and non-outbreak situations (AUC \approx 0.50).

The best results for multivariate outbreak detection were obtained for analyses that combined *NervSy* with *DeadBird* data, which gave similar results to a combination of the three time series (figure 3 and table 2). The results of using *NervSy* combined with *DeadBird* were also better than those obtained with each time series alone. For example, for a specificity set at 0.80, the sensitivity of the detection reached 0.80 with the combined *NervSy* and *DeadBird* series whereas it was 0.67 with *NervSy* and 0.60 with *DeadBird* alone.

DISCUSSION

Our results indicated that the best detection performance was obtained using multivariate syndromic surveillance based on reports of nervous symptoms in horses (*NervSy*) and wild bird mortality (*DeadBird*). To our knowledge, this is the first time that multivariate syndromic surveillance has been implemented for WNV detection. However, when using a univariate detection method, *NervSy* was the best indicator of WNV outbreaks. This is consistent with the number of expected cases during an outbreak compared to the baseline of each time series considered (i.e. high number of case for *NervSy*, moderate number of cases for *DeadBird*, and low number of cases for *DeadHorse*). Indeed, models based only on the *DeadHorse* data resulted in poor detection performance at the regional level because mortality in horses is mainly due to causes other than WNV. However, before ruling on the usefulness of this datasource for WNV surveillance, it would be interesting to test whether an outbreak generates local clusters of deaths in horses that may be used as a signal of a VNW outbreak. However, the quality of geographical information of reported cases are currently insufficient to test this hypothesis.

This is the first time that a real assessment of system performance has been implemented for WNV surveillance. Previous early warning systems developed for WNV only identified risk factors of WNV outbreaks, but did not evaluate the detection performances of those systems (Adlouni et al. 2007; Bellini et al. 2014a; Brown 2012; Chaskopoulou et al. 2013; Gosselin et al. 2005; Rosà et al. 2014; Shuai et al. 2006; Valiakos et al. 2014). Timeliness has occasionally been evaluated but only based on a limited number of real WNV outbreaks, and has not been associated with a further assessment of system performance (Calzolari et al. 2013; Chaintoutis et al. 2014; Eidson et al. 2001; Johnson et al. 2006; Mostashari et al. 2003; Veksler et al. 2009). Only two attempts to assess the sensitivity and specificity of surveillance have been made (Andersson et al. 2014; Leblond et al. 2007) but the parameters of interest were only evaluated based on a limited number of outbreaks, which did not allow any conclusions to be drawn regarding overall system performance.

To assess the surveillance systems and compare them, we simulated baselines and outbreaks using parameters from data observed in Europe (Anonymous 2007; Autorino et al. 2002; Bakonyi et al. 2013; Leblond et al. 2007; Ward et al. 2006). To expand upon this, patterns of outbreaks in other locations

should be tested in order to evaluate the performance of multivariate syndromic surveillance in more varied situations. Particular attention should be paid to patterns of mortality of wild birds, as the dynamics of wild bird mortality during a WNV outbreak have only been poorly investigated in Europe.

The Bayesian approach seems well adapted for multivariate WNV detection and can be used for other diseases. Indeed, Bayesian hypothesis testing is based on two mutually exclusive hypotheses which can be expressed in the syndromic surveillance context as: H_1 , “there is an ongoing outbreak of WNV or of another disease with similar symptoms”, and H_0 , “there is no ongoing outbreak”. It would be theoretically possible to include every possible differential diagnosis for every syndrome (or group of syndromes) considered; however, such a system would be difficult to implement and maintain. It would thus be interesting to first examine the evidence from each time series individually and then together in order to identify which combination of datasets results in the strongest signal. It would be up to the relevant decision maker in a given situation to consider appropriate differential diagnoses and the actions that should be implemented for further investigation.

In our study, we considered three sources of evidence for WNV outbreak detection. Nevertheless, additional information can be utilized with Bayesian approaches, as it is easy to add such information. Then, a next step in the early detection of WNV outbreaks should be to test the efficiency of the method with other data, such as the predicted abundance of mosquitoes (Calistri et al. 2014; Rosà et al. 2014), environmental risk factors (Tran et al. 2014), and probability of introduction (Bessell et al. 2014; Brown et al. 2012).

CONCLUSION

The proposed approach is suitable for performing multivariate syndromic surveillance of WNV outbreaks. Indeed, we found that a multivariate surveillance system using this approach performed better than a univariate approach in detecting WNV outbreaks in southern France. In particular, a combination of data regarding nervous symptoms in horses and wild bird mortality was the most efficient in detecting outbreaks. Such multivariate surveillance systems could be especially useful in serving as early warnings for possible human viral infections, considering that horses and birds are affected by WNV before humans (Kulasekera et al. 2001; Leblond et al. 2007). We propose that this methodology is generally applicable to other diseases for which multiple sources of evidence are available.

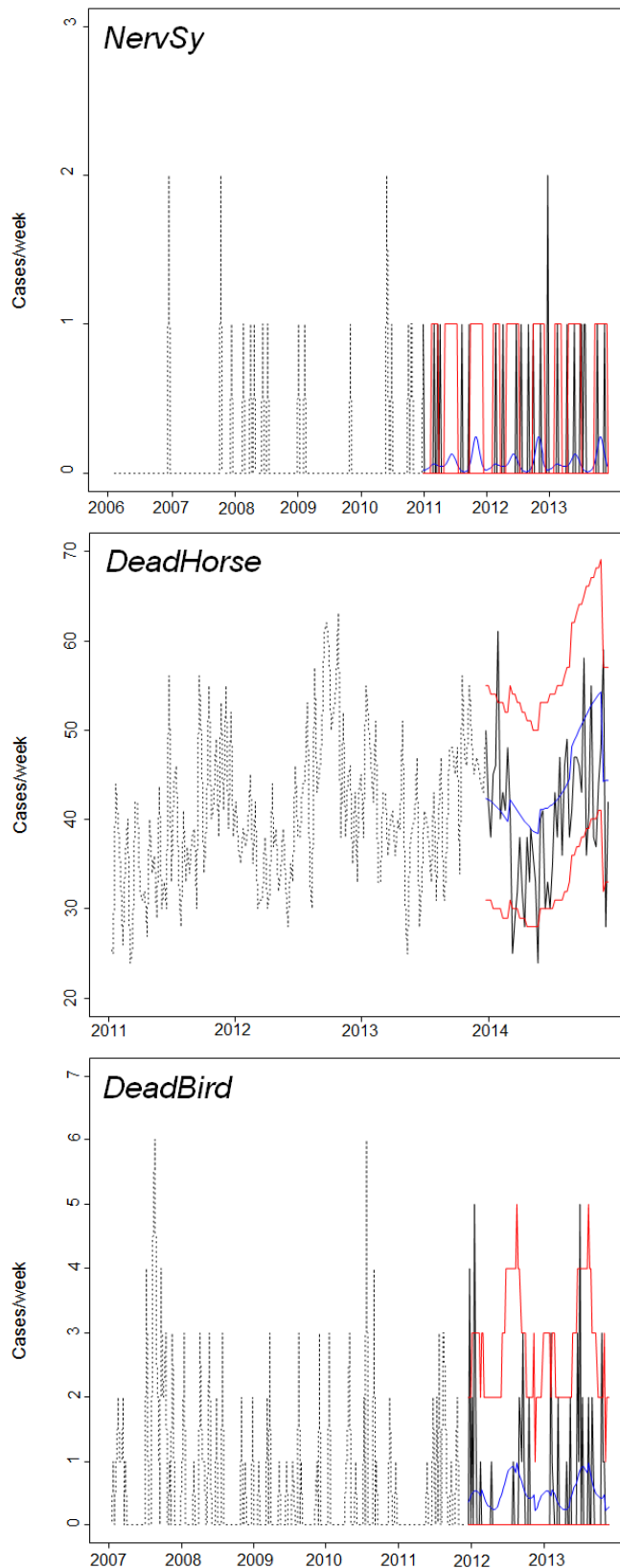


Figure 1: three time series considered. NervSy: number of declaration of nervous syndrome in horses without positive lab result. DeadHorse: number of dead adult horses collected by French fallen stock companies. DeadBird: number of dead wild birds autopsied with values above the 95% confidence interval deleted. Dotted lines = training data, solid black lines = test data, solid blue lines = predicted value, solid red lines = 95% Confidence interval

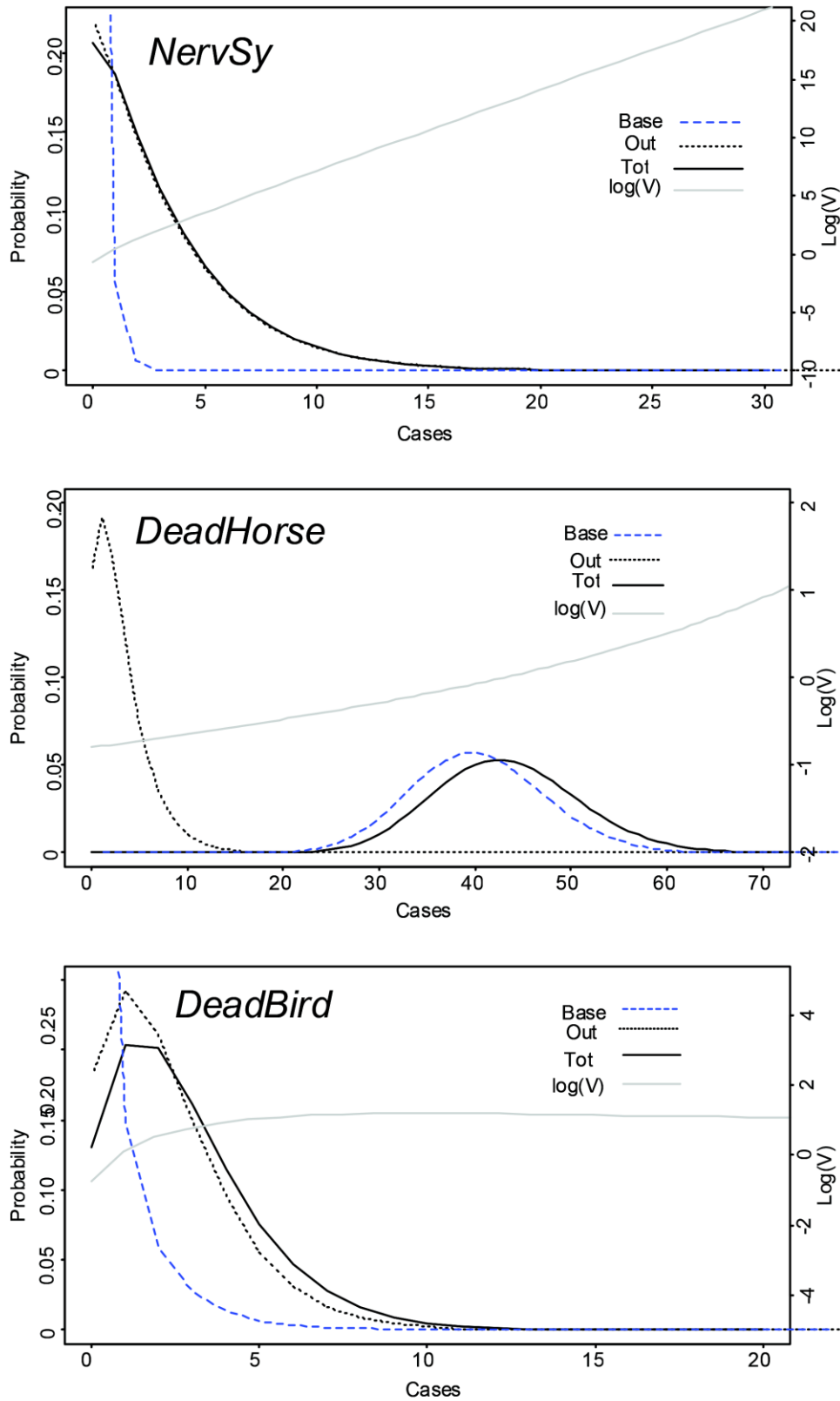


Figure 2: Value of evidence and probabilities of observing n cases during a non-outbreak (Base) and an outbreak (Out) situation. Base= distribution of distribution into the baseline, Out = distribution of cases related to a WNV outbreak, Tot= distribution of cases during an outbreak (Base + Out), $\text{Log}(V) = \log_{10}(p(n | \text{outbreak}) / p(n | \text{baseline}))$. Out was based for NervSy on NB($\mu = 3.12$, $\theta = 1.150$), for DeadHorse on NB($\mu = 3$, $\theta = 2.005$), and for DeadBird on NB($\text{mean} = 2.23$, $\theta = 3.34$).

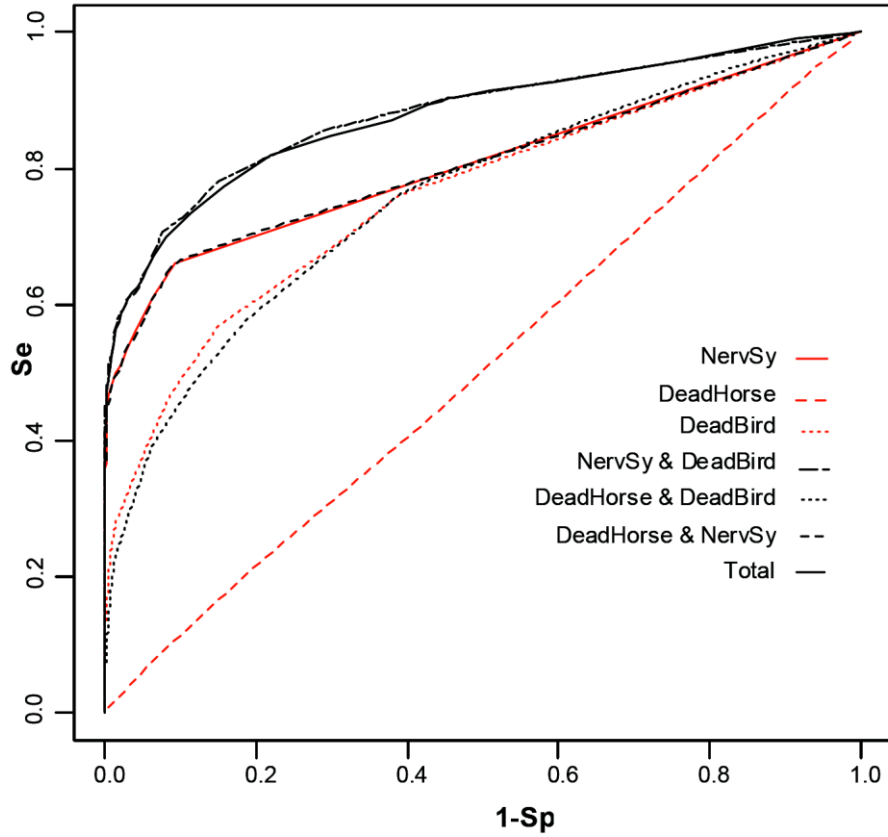


Figure 3: ROC curves for univariate and multivariate outbreak detection using NervSy, DeadHorse and DeadBird.

Negative binomial distribution			AIC	GOF	RMSE _c	RMSE _v
Formulae	theta	mean				
NervSy $\sim \sin(2\pi(t-4)/18.33) + \sin(2\pi t/26.5)$	0.413	0.077	143	0.279	0.30	0.39
DeadHorse $\sim 4 \times (t-4)/52 + t + \sin(2\pi(t-12)/53)$	176	40.3	1063	1.016	7.06	8.57
DeadBird $\sim 4 \times (t-4)/52 + \sin(2\pi t/26.5)$	0.373	0.520	497	0.675	1.03	1.05

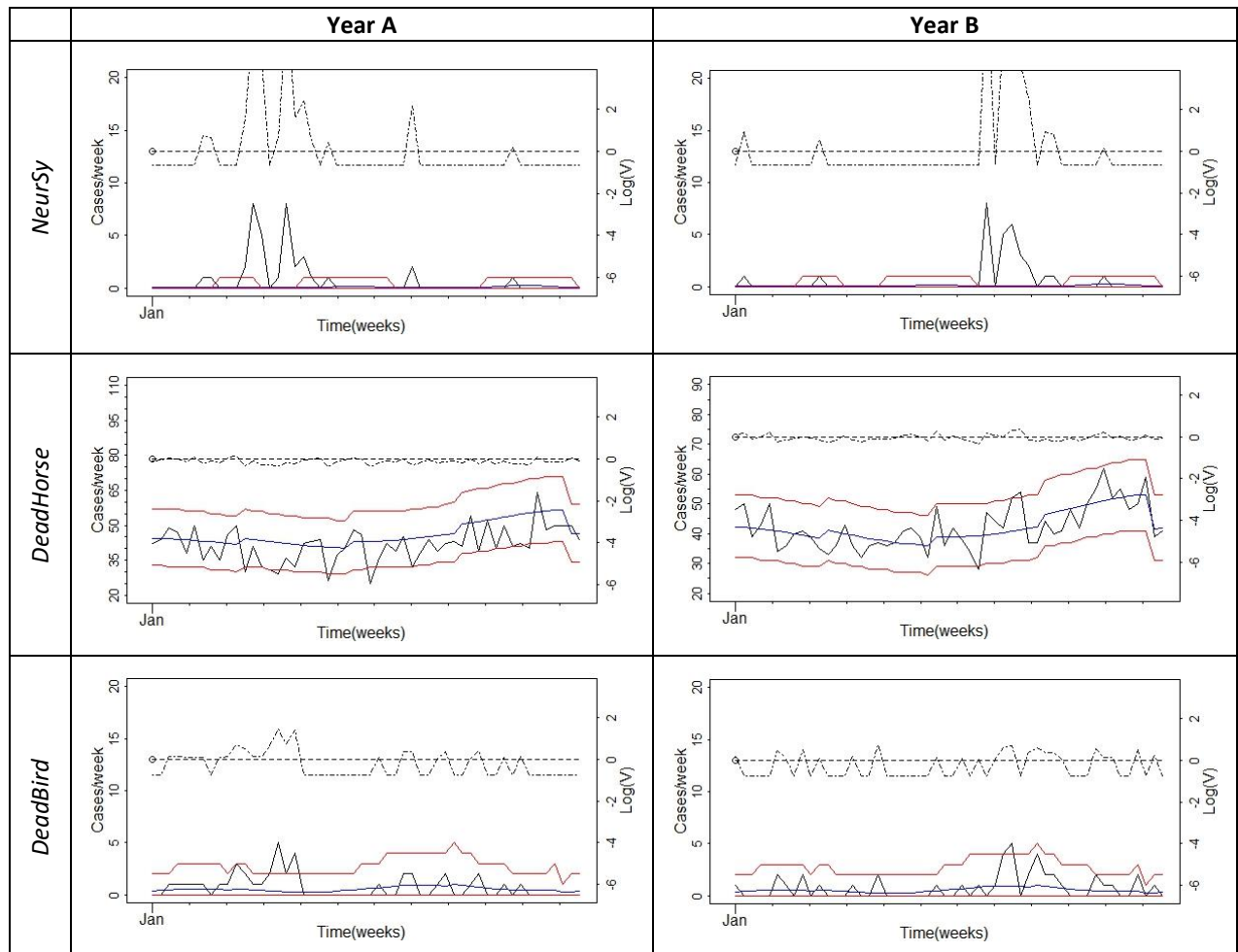
Table 1: Models and models parameters obtained for the three time series.

	NervSy	DeadHorse	DeadBird	NervSy & DeadBird	NervSy & DeadHorse	DeadHorse & DeadBird	Total
AUC	0.80	0.50	0.75	0.87	0.80	0.75	0.87
Standard error	0.0082	0.0097	0.0089	0.0068	0.0081	0.0089	0.0068

Table 2: Area under the ROC curve (AUC) and standard error for univariate and multivariate outbreak detection using NervSy, DeadHorse and DeadBird.

Appendix I:

Supplementary figure 1: Examples of simulated baseline with inserted outbreak and corresponding variation of the value of evidence (V). solid black line = simulated data, solid blue line = predicted value, solid red line = 95% confidence interval, Dotted lines = $\log_{10}(V)$



B. COMBINING QUANTITATIVE RISK ASSESSMENT AND SYNDROMIC SURVEILLANCE

1. Introduction

Risk assessment and syndromic surveillance can both indicate the possibility of a newly introduced epidemic but they do not prove the presence of the disease. Specifically, both approaches provide different risk estimations. Risk assessment identifies the respective probabilities of pathogen entry and establishment according to season and local environment. Syndromic surveillance determines the probability that an outbreak is in progress based on field observations. Combining both approaches will thus result in a posterior probability which should improve confidence in the prediction of an outbreak of a newly introduced VBD. The posterior probability is calculated according to Figure 15.

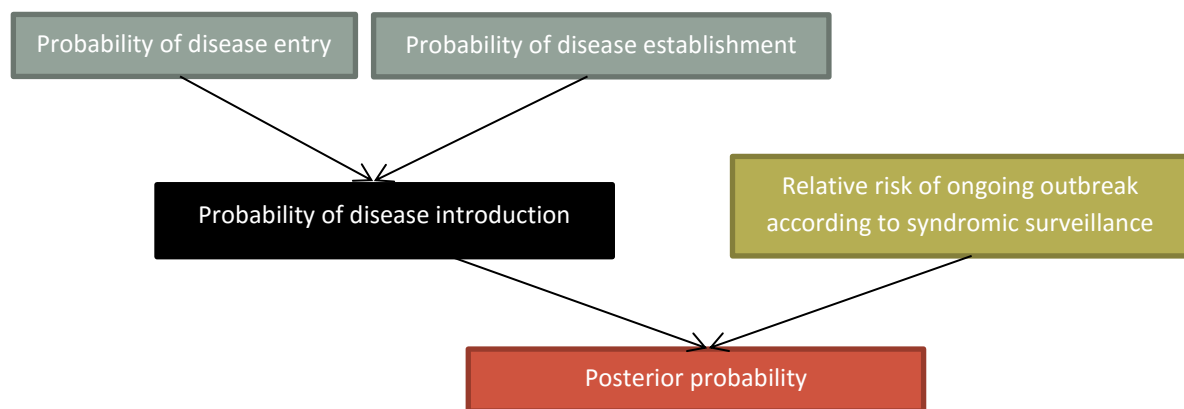


Figure 15: Combining syndromic surveillance and probability of disease introduction to obtain a posterior probability that there is an ongoing outbreak.

In particular, this approach can be useful for exotic diseases with few and unspecific symptoms. Indeed, veterinary practitioners are rarely able to detect such diseases and syndromic surveillance might be useful in enhancing the passive reporting system. However, the incidence of under-reporting is also expected to be high, which compromises the ability of syndromic surveillance to detect disease. Adding prior information on the probability of disease introduction can improve the detection of such diseases.

Among exotic diseases, equine encephalosis (EE) typically causes unspecific symptoms that are quite hard to detect, as illustrated by the example of Israel, where the disease was present for at least 10 years before it was first officially reported (Wescott et al. 2013). In this study we combined syndromic surveillance with information on the probability of disease introduction in order to improve early detection of a potential EE outbreak.

This work will result in a scientific paper in preparation and we only present here preliminary results: Mats Gunnar Andersson*, Egil Andreas Joor Fischer*[†], Céline Faverjon, Jörn Gethmann, Maya Gussmann, Yves

2. Material and Methods

2.1. Bayesian framework

The analyses were performed with weekly data using counts from a grid in France which consisted of 943 cells of 25x25 kilometers each.

To provide a comprehensive and single indicator of risk, we used spatiotemporal extension of the empirical Bayes' approach presented in previous sections (see Chapter III.C and Chapter IV.B). Bayesian hypothesis testing is based on two mutually exclusive hypotheses which can be expressed in the syndromic surveillance context as: H_1 , "there is an ongoing outbreak of EE (or another disease with similar symptoms) in grid cell g during week w " and H_0 , "there is no ongoing outbreak in grid cell g during week w " (Andersson et al. 2014). The relative probability of the two hypotheses can be expressed as a ratio (O_{pri}) that represents our *a priori* belief about the disease status:

$$O_{pri} = \frac{P(H_1)}{P(H_0)}$$

The probabilities of H_1 and H_0 were estimated based on the spatiotemporal quantitative model for the assessment of the probability of EE introduction developed in Chapter II.C. The value of evidence (V) was estimated using syndromic surveillance data. The main symptoms of EE in horses are respiratory and nervous symptoms (Dhama et al. 2014); we thus considered these two sources of data (*NervSy* and *RespSy*, respectively). We supposed both datasets to be independent and describe them in the next section. The corresponding value of evidence (V) was calculated as:

$$V = \frac{P(E_{NervSy} | H_1)}{P(E_{NervSy} | H_0)} \times \frac{P(E_{RespSy} | H_1)}{P(E_{RespSy} | H_0)} = \frac{P(E_{NervSy}, E_{RespSy} | H_1)}{P(E_{NervSy}, E_{RespSy} | H_0)}$$

where $P(E_X | H_1)$ is the probability of observing the number of reported cases of syndrome X in a particular week in a particular grid cell given that H_1 is true, and $P(E_X | H_0)$ is the probability of observing the number of reported cases of syndrome X in a particular week in a particular grid cell given that H_0 is true.

The *a posteriori* belief about the disease's status for each week and each grid cell (O_{post}) takes into account both the probability of disease introduction and the results of syndromic surveillance and was calculated as in (Faverjon et al. 2015a):

$$O_{post} = \frac{P(H_1 | E_{NervSy}, E_{RespSy})}{P(H_0 | E_{NervSy}, E_{RespSy})} = V \times O_{pri}$$

2.2. Simulated EE outbreaks

Two outbreaks were simulated to test our model: one in Aquitaine, and another in Normandy. These regions were selected due to their large equine population. No spatiotemporal description of EE outbreaks is currently available, so we used other, indirect, information to estimate the number of expected cases and the spatiotemporal evolution of an EE outbreak.

First, we assumed that 80% of the susceptible population would be infected, based on the estimated prevalence of EE in the initial serological data from Israel in 2001 (Wescott et al. 2013). To obtain the number of horses showing clinical signs for EE, we supposed that around 40% of the infected horses would show clinical signs and, of these, 90% would be respiratory symptoms and 10% neurological symptoms (Aharonson-Raz et al, 2011). We then presumed that 25% of respiratory cases and 50% of neurological cases would be declared to RESPE. In the end, we estimated that, during an EE outbreak, the percentage of horses with respiratory or nervous symptoms due to EE and declared to RESPE was 7% and 1.6%, respectively, of the general equine population located in an area.

To obtain the distribution of horses that showed clinical signs per week, we used data collected during African Horse Sickness (AHS) outbreaks, because EE and AHS are very similar and share the same transmission patterns (Dhama et al. 2014; Lord et al. 2002). Specifically, we used data from outbreaks in Western Cape in 1999, 2004, and 2011 (Anonymous 2011; Sinclair et al. 2006). These outbreaks occurred in an area of South Africa that is under surveillance designed to act as an early warning system, but which lacks systematic vaccination campaigns performed to protect the free zone. Nevertheless, the number of vaccinated horses is considered to be high in this area (Sinclair et al. 2006), which probably reduced the size of the outbreaks compared to a situation in which all horses are susceptible.

To estimate the spatial propagation of EE for the outbreak in “Aquitaine”, we used the first six weeks of location data from a BTV-1 outbreak in southern France, as both viruses share the same vector, *Culicoides* (Venter et al. 2002). For the outbreak in “Normandy”, we assumed that the spread of an EE outbreak would be similar in velocity to that observed during the last BTV-8 outbreak in France: from an index case, the outbreak gained around 10 kilometers per week (Pioz et al. 2008), with 50% of new cases occurring less than 5 km from the closest infected animal, and 95% within a radius of 31 km of the closest infected animal (Hendrickx et al. 2008). The distribution of cases has been made according to local equine population using data provided by IFCE-SIRE.

The spatiotemporal progression of the outbreak in Aquitaine is presented in Figure 16, while the outbreak in Normandy is available in Appendix 11.

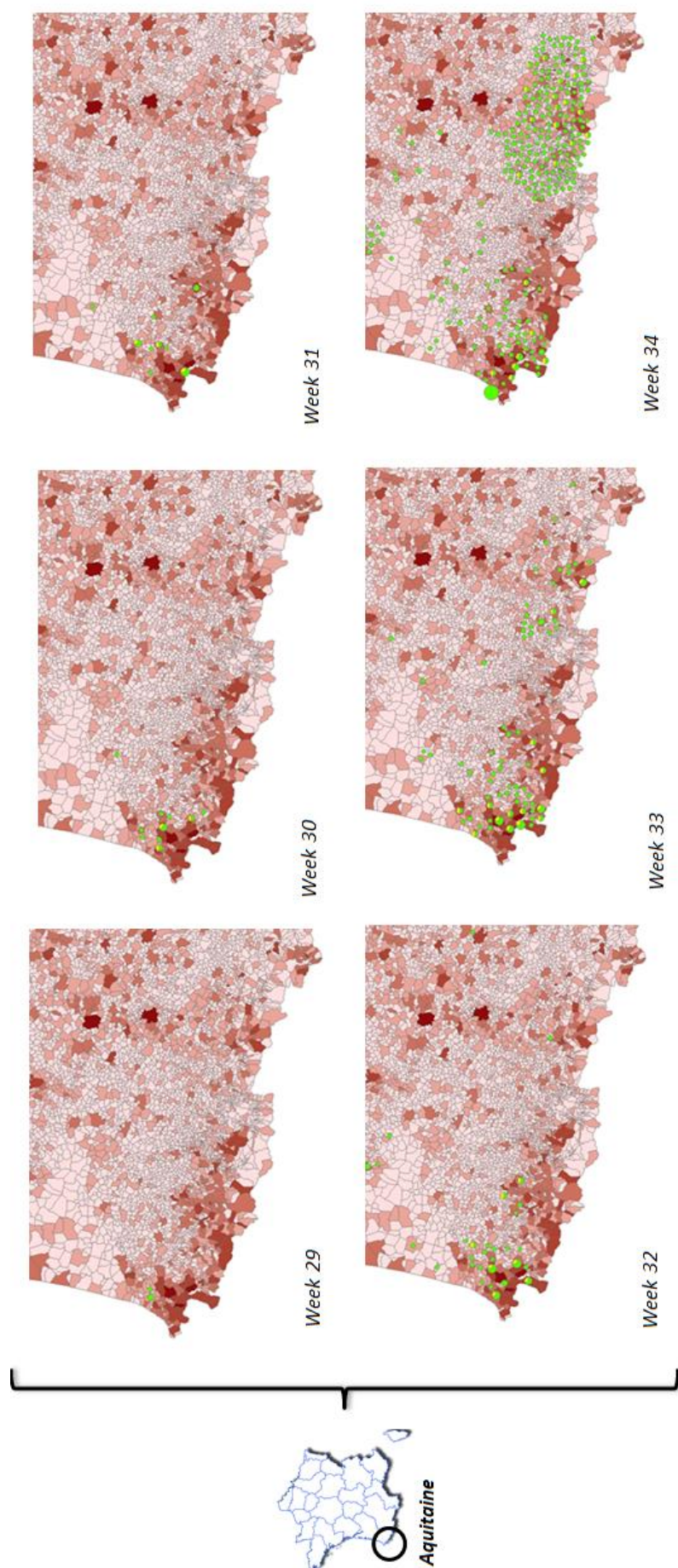


Figure 16: EE simulated outbreak in Aquitaine. The outbreak starts mid July 2012 and lengths 6 weeks (from week 29 to week 34). Yellow circles = nervous cases (total 59), green circles = respiratory cases (total 536)

2.3. Probability of EE introduction

2.3.1. Data source

To define our *a priori* belief about the disease status (O_{pri}), we reused the spatiotemporal quantitative model that was developed in Chapter II.C to assess EE introduction. The model served as a basis for estimations of the probability of disease entry and the probability of disease establishment. The spatiotemporal model of EE introduction uses monthly data at the regional scale (22 regions) for three consecutive years (2010, 2011, and 2012).

2.3.2. Data transformation

To transform the monthly and regional data into weekly, grid-based data, we assumed that the monthly regional probabilities of virus introduction were uniformly distributed in space and time within a given region and for a given month. The weekly probability of virus release in a grid cell, $P(intro_{wg})$, is thus defined for both routes of introduction as:

$$P(intro_{wg}) = 1 - (1 - P(intro_{km}))^{\frac{1}{week_m \cdot grid_k}}$$

where w belongs to month m , g belongs to region k , $week_m$ equals the number of weeks in month m , $grid_k$ equals the number of grid cells in region k , and $P(intro_{km})$ is the monthly regional probability of virus introduction via an infectious host or vector.

The probability of introduction sometimes equaled zero, due to a lack of importation or to a null probability of virus circulation in exporting countries. When combining this zero probability with the syndromic surveillance model, the final output was thus zero as well. We adopted a conservative approach and assumed that there is always a slight probability of virus introduction and transmission; we thus chose to replace these zero probabilities of introduction with a value equal to 10% of the lowest calculated probability of introduction.

2.4. Syndromic surveillance of EE

2.4.1. Data sources

To define the value of evidence (V), two types of syndromes were considered: nervous symptoms and respiratory symptoms in horses. Both sets of related data are collected by RESPE and have already been used and presented in Chapter III.C. The datasets without positive laboratory diagnoses were used to obtain an outbreak-free dataset. Syndromic surveillance data collected by RESPE are available on a daily basis and at the municipal level from 2006 to 2013.

2.4.2. Spatiotemporal detection

The algorithm used for spatiotemporal detection of clusters of cases was developed and implemented by Gunnar Andersson (SVA, Sweden) within the framework of the EMIDA-VICE European project (Andersson et al. 2015). Briefly the seasonal variation of each syndrome/symptom, under outbreak-free conditions was first modelled for France as a whole. The probability distribution of the number of reported syndromes/symptoms around each grid-cell was subsequently modelled using the expected number of cases per host, and local host density as input. The host population horse population around each grid cell was based on data at commune level provided by IFCE-SIRE (IFCE - les Haras nationaux 2011)).

2.5. Presentation of the concept and first feedback

To evaluate and receive feedback on the proposed approach, the concept was presented during a two-day workshop organized by the EMIDA-VICE project in Paris in February 2015. Fourteen people not included in the VICE project from eight countries and one person representing EFSA were present. The participants were from research centers, state agencies, or veterinary services, and were all involved in surveillance (data collection and management), risk assessment, and/or risk management.

In addition to the general approach, we also presented the EE example. Practical exercises were provided to the participants to facilitate their understanding of the concept. Their comments and suggestions on the concept and method were collected.

3. Results

3.1. Multivariate syndromic surveillance

Models that made use of both syndromes (neurological and nervous) discriminated better between outbreak and non-outbreak situations than those that considered symptoms separately (see Figure 17).

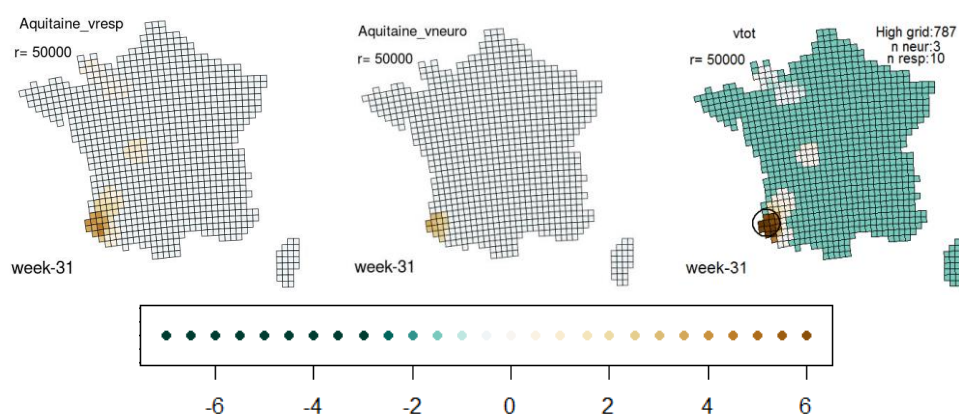


Figure 17: Value of evidence in France at week 31 in year 2012 from the simulated outbreak in Aquitaine, using a detection radius of 50 km. Map on the left = respiratory symptoms, middle map = nervous symptoms, map on the right = multivariate syndromic surveillance (Andersson et al. 2015)

3.2. *Combining Risk assessment*

The combination of the probability of disease introduction with syndromic surveillance results is presented in Figure 18. The results for the outbreak in Normandy are presented in Appendix 12.

In both cases, combining prior information on the probability of disease introduction (O_{pri}) with a multivariate syndromic surveillance system (V) allowed us to i) highlight a small number of cases occurring in an area at high probability for disease introduction, and ii) diminish the significance of a large number of cases occurring in an area at low probability for disease introduction.

3.3. *Workshop output*

The workshop organized by the EMIDA-VICE project presented our approach and the EE example to fourteen participants from eight European countries. The main conclusion of the workshop was that the approach was of great interest for combining different risk assessments. The participants highlighted the wide range of opportunities offered by this tool, but were of the opinion that this approach would mainly be used to increase awareness, rather than to implement active surveillance. Indeed, for the participants, the approach was helpful for demonstration and should facilitate risk communication (in particular when different scenarios are tested). The main constraint to the implementation of this approach was, according to the participants, the current lack of relevant and accurate data.

Despite these advantages, the approach also appeared quite complicated, and the participants advocated for the automatic generation of output, a user-friendly interface, a list of requirements to run the system, and an easy-to-use manual. Other criticisms were that the assumptions were not sufficiently explicit and the uncertainty was not displayed, which complicated the understanding and interpretation of results. The participants also advocated for the development of sensitivity and cost-efficiency analyses for this approach.

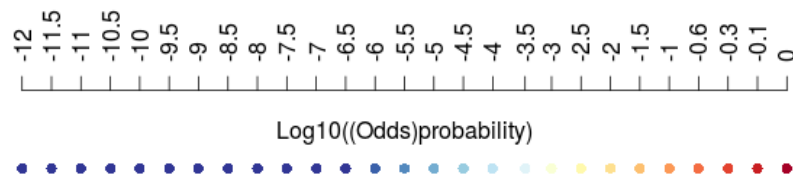
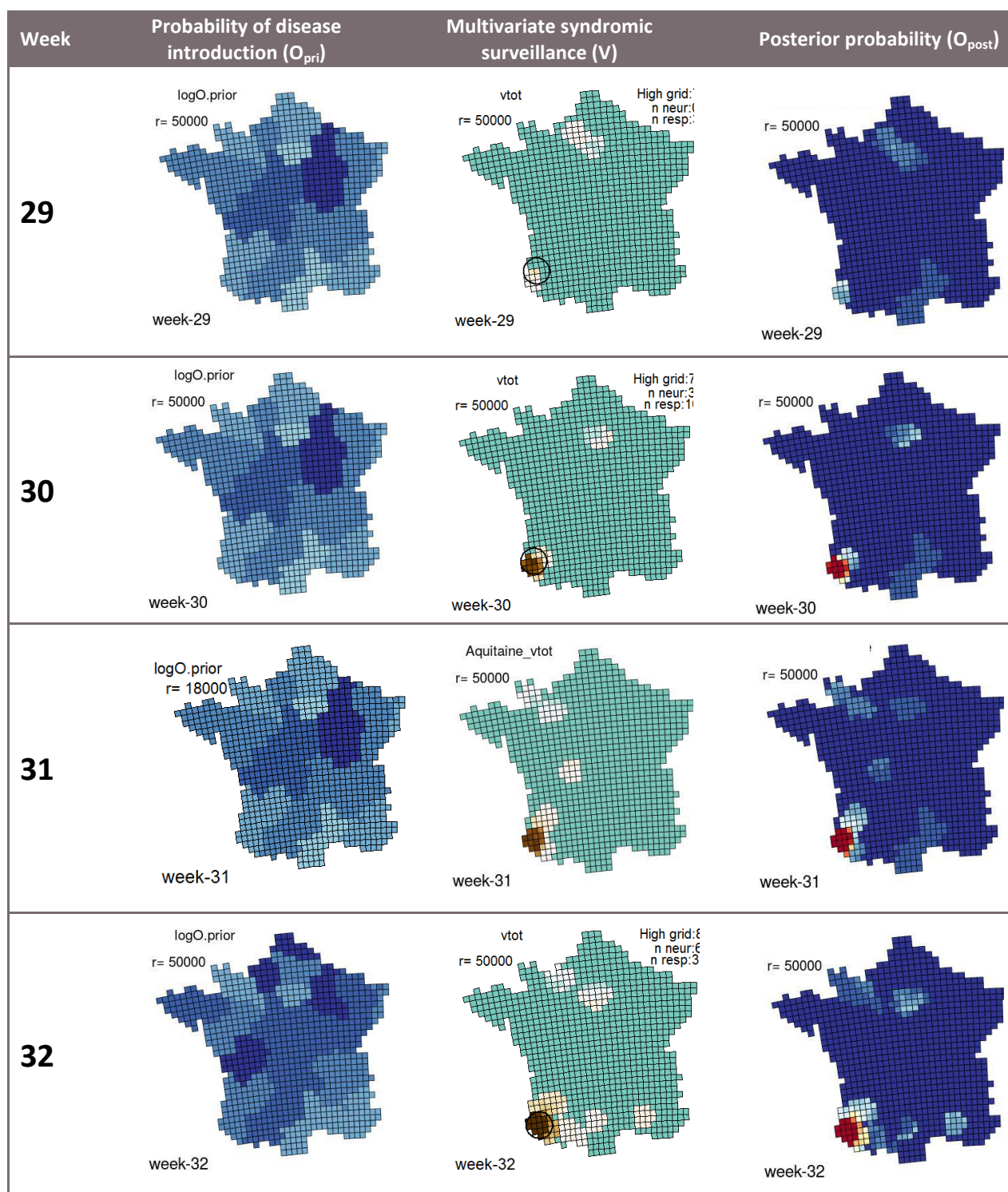


Figure 18: Posterior probability of EE outbreak in France in year 2012 during the simulated outbreak in Aquitaine from week 29 to week 32. Radius for detection is 50 km (Andersson et al. 2015)

4. Discussion and Conclusion

The results presented are only preliminary but they provide a good example of the concept. In this example, the approach was applied to detect a known disease (equine encephalosis), but it also presents multiple other opportunities. Indeed, any kind of prior knowledge can be used and thus the approach could also be implemented for diseases about which not a great deal is known *a priori*. Moreover, it can be used not only in the context of risk assessment, but also in that of risk management. For example, the approach can be used to test different risk scenarios and increase awareness, and thus support decision-making. It is particularly adapted for use in decision-making because it can be easily combined with decision theory, as presented in (Andersson et al. 2014).

The feedback from the EMIDA-VICE workshop emphasized the potential of our approach but also highlighted a limitation to its practical implementation: the lack of relevant and accurate data. . Indeed, the approach requires a large amount of data (regarding, e.g., import of animals, temperature, vector abundances, declarations from veterinarians), which are currently very rarely available all together. At the European scale, at the time of writing, there is no agreement on data format or on the definition of terms, especially in syndromic surveillance. Current initiatives to collect and analyze data are thus rarely transferable between EU countries and the results of these programs are not comparable to each other. Further efforts should be made to homogenize data collection and the definition of terms at the European level in order to facilitate the widespread use of the results of the different existing surveillance systems. However, even if the lack of data is a concern, Bayesian framework allows us to deal with missing data by combining expert opinion with data driven models. The lack of relevant and accurate data is thus not a fundamental limitation of the approach.

In this study, the example of equine encephalosis served mainly to illustrate our concept. That is why we only considered a simple approach for outbreak detection. However, this approach had some limits. In particular, we assumed that the population at risk remained stable over time, and we detected clusters of cases based on the size of this baseline population. This assumption can be questioned, particularly regarding horses, which travel frequently and for which no accurate data exist on numbers and geographical localization. Other methods of cluster detection should be tested (for a summary of the main statistical methods available for testing clusters in space, time and space, and time*space (interaction), see the review by T. Carpenter (Carpenter 2001)) and other detection parameters should be also considered to optimize outbreak detection. In particular regarding detection parameters, radius' size on outbreak detection, temporal units of detection and potential interactions among cluster parameters (Olson et al. 2005; Ozonoff et al. 2007). From the perspective of a real outbreak detection system, all these parameters should be carefully investigated and their appropriate values determined, in collaboration with decision makers, according to the objective of the surveillance and to the disease considered. Moreover, in this example we used respiratory and nervous symptoms observed in horses, but it could be interesting to evaluate other symptoms, such as spontaneous abortions, which are often

present during EE outbreaks (Dhama et al. 2014). Testing combinations of other symptoms would thus be an important part of optimizing outbreak detection.

Using a Bayesian approach for combining evidence is not new, but it is still rarely implemented in disease surveillance. To our knowledge, there are only a few examples in the literature of this kind of data combination in disease surveillance. We can cite the work of Gustafson and colleagues (Gustafson et al. 2010), which proposed using the likelihood ratio to combine expert opinion with surveillance data in surveys of viral hemorrhagic septicemia, but the work of Burkom and colleagues (Burkom et al. 2011) is perhaps closer to our concept (i.e. combining health surveillance data and environmental sensors for monitoring waterborne diseases). Bayesian approaches are thus still new in disease surveillance and offer great opportunities, but, as highlighted during our workshop, further work must also be conducted. In particular, a full assessment of the tool is fundamental in order to know more about its sensitivity and reliability. Moreover, further developments should focus on making our approach more comprehensive and user-friendly, or it will likely not be adopted by decision-makers and other participants in disease surveillance.

C. DISCUSSION AND CONCLUSION

1. Discussion

The complementary approaches used in this chapter (i.e. multivariate syndromic surveillance and the combination of probability of introduction with syndromic surveillance) were both effective in providing a better estimation of the probability of an ongoing outbreak compared to estimates based on a single probability estimation. Similar methods are nonetheless rarely implemented in veterinary public health. Multivariate syndromic surveillance is still a new approach and, at the time of writing, it has been reported from only one initiative in this field (Höle et al. 2007), which used multivariate CUSUMs for rabies surveillance. Instead, combining syndromic surveillance with quantitative risk assessment is a completely innovative concept and, to date, no similar approach has been presented in the veterinary public health literature.

Multivariate risk assessment is heuristically better than univariate risk assessment. However, there are a number of practical issues that make multivariate assessment difficult to implement. For example, the multivariate syndromic surveillance systems developed in this chapter assumed that the data sources were independent and that changes during an outbreak occurred simultaneously in all data considered. These assumptions were simple but justifiable with WNV surveillance that used nervous symptoms in horses, mortality in wild birds, and mortality in horses. However, adapting the approach to another disease may require different assumptions and thus models of greater complexity, in order to deal with interactions between risk sources (Frisén et al. 2010). This is not a fundamental limitation of the approach, but further complications in model implementation may place limits on its practical use. Similarly, we combined syndromic surveillance and risk assessment assuming that both were *a priori* independent, as they do not utilize the same data sources. However, this was not completely true, and in reality the simultaneous use of syndromic surveillance and risk assessment is not so simple. As an example, risk assessment can be used to enhance veterinary practitioners' awareness through the production of risk maps, as has already been proposed for some endemic VBDs (e.g., surveillance of West Nile virus in California (Brown 2012) or tick-borne diseases in Europe (Beugnet et al. 2009)). Enhancing awareness in this way plays an important role in improving the early detection of disease. However, syndromic surveillance is closely dependent on practitioners' awareness, and increased awareness can result in an increasing number of (potentially spurious) declarations, and thus more false alarms. When combining risk sources, then, one must be careful to take into account any possible correlation between them.

2. Conclusion

Using a combination of risk estimations provided a better estimation that an outbreak might be ongoing. However, the interactions between risk estimations must be carefully considered, as they may have an

impact on the complexity and the performance of the surveillance system. This is of particular concern regarding the practical implementation of such a combined surveillance system.

CHAPTER V: DISCUSSION

The objective of this work was to address the challenges of establishing early warning systems for VBDs in horses by using quantitative risk assessments and syndromic surveillance, alone and in concert. All the methods developed in this work provided estimates of the probability of a VBD outbreak in horses and may help to address the challenge of VBDs surveillance in horses. However, many questions were raised in the implementation of these approaches (i.e. quantitative risk assessments and/or syndromic surveillance).

In this chapter, we start by presenting a brief summary of the main results obtained regarding the specific case of VBDs in horses, and we balance these against our initial objective and work already conducted on this topic. Then, from a wider perspective, we address how our approaches may support decision making and how they relate to demonstrating the absence of disease. Finally, we discuss the practical implementation of such methods.

1. Early warning system of VBDs in horses

1.1. *Quantitative risk assessment*

In this work, we showed that combining the probabilities of entry and establishment and taking into account spatiotemporal aspects in a quantitative assessment of the probability of import were both essential in obtaining a complete view of the risk posed by a vector-borne pathogen. This is consistent with the close link between VBDs and their climate and environment. By combining two routes of pathogen entry, we were able to better understand the risk posed by a pathogen to the equine industry. Additionally, although quantitative risk assessment is not a new method for assessing the probability of outbreak occurrence, spatiotemporal analyses and combinations of probabilities and of entry routes are still rarely implemented in animal health. Our work thus advocates for the wider use of these kinds of approaches in order to obtain a more detailed and complete picture of the risk. It could also be interesting to apply this method to studies of other pathogens. In particular, assessments of the probability of bluetongue virus introduction could be easily implemented, as this virus is *Culicoides*-borne and is similar to the viruses responsible for African horse sickness and equine encephalosis. This part of the work also highlighted some limits specific to the equine industry. In particular, the available data were sometimes inaccurate because the equine industry is fragmented and poorly tracked, and few databases exist. When data do exist, they are rarely centralized, which presents problems regarding accessibility and standardization. Further work should be conducted to improve the quality of data on horses. It would be particularly helpful to reinforce the tracking of horse movements and to continue with efforts to identify the location of horses using awareness campaigns or field surveys.

1.2. Syndromic surveillance

Syndromic surveillance of nervous symptoms in horses using an alarm threshold that was a multiple of the standard error of prediction was able to detect early signals of a WNV outbreak in French horses. However, our study represents the first time that a full assessment of this system's performance has been carried out. Using a simple Bayesian approach, we were able to provide a quantitative evaluation of the probability of an outbreak and generate an easy-to-interpret output that was simple to combine with other epidemiological knowledge, such as disease seasonality. This is especially important for the surveillance of VBDs, which are closely related to season. The output can be also combined with socio-economic information in order to facilitate a more transparent and evidence-based decision-making process. To date, syndromic surveillance has rarely been implemented in animal health and even less often for diseases in horses, a deficiency that our work seeks to address. In addition to the diseases studied here, syndromic surveillance could also be used to monitor for the presence of other exotic diseases that induce nervous symptoms in horses, such as Eastern and Western equine encephalitis, Venezuelan equine encephalitis, or Japanese encephalitis, which are of serious concern for public health. It would also be interesting to explore other symptoms occurring in horses. Here, we considered only nervous and respiratory symptoms, but future surveillance systems could also use data on abortions, for example, to detect equine arteritis virus, another disease of interest for the equine industry. In the present work, we did not quantitatively compare the respective performances of both proposed approaches (i.e., classical approach with an alarm threshold that was a multiple of the standard error of prediction and Bayesian approach) in outbreak detection, but this step would also be important for future work.

1.3. Combining risk

To improve confidence in the prediction of an ongoing outbreak, we combined different sources of risk assessment, first by using multivariate syndromic surveillance. This approach provided better detection performance than univariate syndromic surveillance, but it is rarely implemented in veterinary public health. It could however be applied in a wide range of situations. Considering other VBDs that affect horse populations, it could be also used, for example, to monitor for Japanese encephalitis, by combining surveillance of nervous symptoms in horses with that of reproductive diseases in swine. As a further step in risk combination, we also combined assessments of the probability of introduction with syndromic surveillance. Such integrated surveillance systems already exist in animal health, in particular for WNV surveillance. For example, the system implemented in California provides updated risk maps in which all information is gathered together through a scoring system (Brown 2012); similar systems also exist in Canada (Gosselin et al. 2005), Italy (Bellini et al. 2014a), and Greece (Chaskopoulou et al. 2013). However, unlike our approach, these integrated surveillance systems do not provide a single and quantitative output that combines all the collected information. Our approach thus represents a promising way to build a surveillance system that can quantitatively combine several estimations of risk in a single risk indicator.

This method could be of particular use for the surveillance of VBDs, as it can include information on seasonality and environmental risk factors, which are both fundamental parameters of VBD biology. In addition, combining syndromic surveillance with other epidemiological information is a way to improve the specificity of detection and could thus strengthen the surveillance of VBDs. This approach could also be applied to various other exotic diseases, such as equine encephalitis or exotic strains of bluetongue virus. However, the study presented here of the combination of syndromic surveillance with risk analysis was only preliminary, and further work should be conducted, especially to investigate tradeoffs between system performance and system complexity.

2. Support decision-making

The present work was primarily interested in quantitatively combining different sources of risk (i.e. routes of pathogen entry, probabilities of entry and establishment, syndromic surveillance data sources, syndromic surveillance and risk assessment). Apart from the obvious benefits of producing more reliable information, such an approach is also a good way to further support decision making in veterinary public health, as shown, for example, during the workshop described in Chapter IV.B

At the individual level, decision making is the cognitive process that leads to the selection of a course of action and ends with a final choice. Decision making is a part of risk management (OIE 2010) and is based on a complex combination of rational analysis and subjective opinion (Damasio 1995, Slovic et al. 2005). In public health, there is a consensus that decision making should be strictly rational and based on a combination of scientific evidence, available resources, and context (Brownson et al. 2010). The concept of evidence-based practice dates from 1971 (Cochrane 1999). It was initially developed for clinical medicine but the evidence-based philosophy has now also been adopted in public health. Evidence-based public health (EBPH) can be defined as a 'public health endeavour in which there is an informed, explicit, and judicious use of evidence that has been derived from any of a variety of science and social science research and assessment methods' (Rychetnik et al. 2004). EBPH and the related evidence-based veterinary public health (EBVPH) are still in their early stages (Latham et al. 2013), although guidelines and tools have recently emerged from international agencies (see, for example, the European Centre for Disease Prevention and Control (ECDC 2011) and the Center for Disease Control and Prevention (Jacobs et al. 2012)). The fact remains, though, that in the real world, decisions in public health are rarely made using empirical-analytical models, but are often conjectures based on crises, hot issues, short-term opportunities, concerns of organized interest groups, political and practical judgments, and public concern (Brownson et al. 2010, Head 2010, Rutherford et al. 2010, Sanderson 2002).

Several factors explain the limited use of EBPH, such as the lack of understanding of evidence-based methodologies by policy makers and the lack of relevant data and formalized systems (Brownson and Jones 2009, Dobbins et al. 2004, Latham et al. 2013, Lomas 1997, Rutherford et al. 2010), but also the absence of agreement on how to interpret and compare different types of evidence (Dobbins et al. 2007,

Latham et al. 2013, Tannahill 2008). Indeed, decision makers in EBPH and EBVPH often have to combine different types of evidence and the decision process can be very complex (ECDC 2011). This problem has become particularly acute because the volume of information has increased exponentially and diverse sources of data are rarely presented together, making gathering, synthesizing, and interpreting them an increasingly challenging task (Rutherford et al. 2010).

By providing decision makers with a single risk indicator that synthesizes different risk assessments about the presence of an outbreak, we thus hope to promote the development of EBVPH. Our approach allowed us to easily combine in a single quantitative risk indicator as much evidence as needed. For example, in Chapter IV we combined different syndromic surveillance data sources (i.e. multivariate syndromic surveillance), as well as considering syndromic surveillance together with the probability of pathogen entry. However, more complex models can be built that involve other risk estimations. The present work is thus especially relevant for efforts to facilitate decision making. In particular, the work conducted in Chapter III.C illustrates how risk analysis can fit easily with decision theory and cost-efficiency analysis, yet another way to help decision makers and to promote a more rational decision-making process in veterinary public health. Nevertheless, further work remains to be done on determining the cost of diseases and disease surveillance (Babo Martins and Rushton 2014).

3. Demonstrate freedom of disease

Although risk indicators never prove the presence of a disease, they suggest the likelihood of its existence in a particular place. This is of course particularly true for indicators of combined risks. It can thus be problematic for decision makers to prove freedom from disease in high-risk areas.

Demonstrating that a country is free from a disease is a complex issue, and it is rare to prove absolute freedom from disease (with the exception of highly contagious diseases for which, when there is no case, there is also no disease). According to the OIE Terrestrial Animal Health Code (OIE 2014), a free zone is defined as a zone in which the absence of the disease under consideration has been demonstrated by the requirements given in the Code for free status. Given that surveillance does not detect any infected animal (S^-), the probability of freedom (D^-) is estimated, with the final result that, if the disease is nevertheless present, its prevalence will be lower than a set threshold by a certain level of confidence.

Based on (Martin et al. 2007a; Martin et al. 2007b), the FAO manual on risk-based surveillance (FAO 2014) proposed a method for estimating the probability of freedom (P_{free}) using Bayes' theorem and accumulated historical information. P_{free} is calculated as:

$$P_{free} = P(D^- | S^-) = \frac{(1 - Prior) \times Sp}{(1 - Prior) \times Sp + Prior \times (1 - Se)}$$

where Sp and Se are the specificity and sensitivity of the surveillance system and $Prior$ is the prior probability that the country is infected.

However, the resulting probability of freedom is expected to be quite low. An accumulation of historical information about the probability of freedom over time is thus used to build a better estimate of this probability. $Pfree$ for the previous time period is calculated and used as the prior for the current time period of evaluation. In other words, $Pfree_t$ calculated at time period t can be estimated with:

$$Pfree_t = \frac{(1 - Prior_t) \times Sp}{(1 - Prior_t) \times Sp + Prior_t \times (1 - Se)}$$

where $Prior_t$ is the resulting probability of two non-exclusive states: “the country was not free from the disease to begin with” and “the country became infected during the time period considered”. $Prior_t$ is thus calculated as:

$$Prior_t = Pfree_{(t-1)} + Pintro - Pfree_{(t-1)} \times Pintro$$

Applying such an approach to the surveillance systems presented in this work is thus a way to calculate the probability of freedom using evidence that suggests the presence of the disease. In this case, Se and Sp are the sensitivity and specificity of the syndromic surveillance system and $Pintro$ is the probability of disease introduction. It would also be possible to use either the combination of syndromic surveillance and risk analysis or each approach individually to demonstrate freedom from disease. Of course, such an approach would only be relevant for an exotic infection that spreads slowly and causes few symptoms. For example, in the work conducted here, this method could be used for equine encephalosis but not for African horse sickness.

4. Practical implementation of integrated surveillance systems

Providing decision makers with a single risk indicator that synthesizes the different risks related to an outbreak could potential be very useful, but also results in a complex surveillance system which can be complicated to put into practice.

Firstly, such a complex surveillance system requires a large amount of data, which is not always easy to obtain (e.g., animal movements, vector abundance, host geographical localization, declarations from field workers). In addition, even when data are available, they are often formatted in many different ways, especially at the European level (e.g., different definitions, different geographical units). This is of concern for the use of the data (i.e. a lengthy initial step of preprocessing is needed) but also for sharing the results with partners involved in disease surveillance. However, sharing surveillance results is particularly important for the European Union, where movements of animals and humans are not (or only poorly) tracked within the community. Disease surveillance in the EU therefore depends in part on reliable and up-to-date data-sharing among member countries. The lack of high-quality and standardized data

constitutes a problem not only for the initial implementation of a system, but also for its maintenance. It is especially a concern for syndromic surveillance systems in which reporting has to be continuously stimulated in order to obtain data. Improvements in data collection may be encouraged, however, by proof that the data are important for disease surveillance. Although initial efforts, such as the present work, might be complicated to implement due to the lack of good data available, they can also be used to promote the need for better data collection.

Secondly, the complexity of a surveillance system may result in a reluctance to trust it on the part of decision makers, given the large amount of information and uncertainties involved. The Bayesian approach used in the present work partially solved this concern. This approach is transparent and offers an explicit separation of assumptions, scientific evidence, and criteria for decisions. However, further work should be conducted to ensure its proper communication and acceptance. Indeed, a complex system that combines different risk estimations requires several underlying assumptions, and, before any practical implementation of such a system, it would be necessary to thoroughly explain how to use it and interpret the output. In addition, future work must include a full assessment of system performance, the only guarantee of the reliability of the surveillance outputs. However, it remains challenging to evaluate multivariate surveillance methods due to the several dimensions and complex time relations involved (Frisén et al. 2010). One potential approach that could help would be the use of Monte Carlo simulations, as proposed by Frisé and colleagues (Frisén et al. 2010).

Many different issues remain to be solved before such complex surveillance systems can be broadly applied in veterinary public health. However, the work conducted here shows that these approaches have huge potential and constitutes a promising initial step. Future work should focus on thorough assessments of system performance and effective communication to interested parties. However, as more data become available online and access to data improves, such systems will play an invaluable role in future disease monitoring efforts.

5. Conclusion

The present work proposed to improve the surveillance of vector-borne diseases in horses through different approaches that assessed the probability of occurrence of a newly introduced epidemic. First, we developed a model of quantitative risk assessment to improve estimates of the probability of pathogen introduction. In particular, we performed a spatiotemporal analysis, simultaneously analyzed two routes of virus entry, and also combined the probability of virus entry with the probability of virus establishment. Second, we implemented and assessed syndromic surveillance systems based on two approaches: a classical approach with an alarm threshold based on the standard error of the prediction, and a Bayesian approach based on a likelihood ratio. The Bayesian approach was especially useful as it provided a quantitative assessment of the syndromic surveillance output and was able to combine different information. We therefore also used this approach to combine various sources of risk estimation in order

to improve the assesment of the probability of occurrence of a newly introduced epidemic. We performed multivariate syndromic surveillance and also combined quantitative risk assessment with syndromic surveillance. Approaches that quantitatively combined evidence provided promising results. This work, based on risk estimations, strengthens the surveillance of VBDs in horses and has potential in supporting decision making. In the end, we hope to encourage the improvement of data collection and data sharing, stimulate the implementation of a full assessment of complex surveillance system performance, especially in terms of cost-efficiency, and promote the adoption of the approach by decision makers and other parties involved in disease surveillance through effective communication and training.

APPENDICES

Appendix 1: Model calculation for PW-host. Details of calculation regarding the AHSV introduction via the import of an infectious host.

Every calculus below is made for an equine from species i imported from an area j to the free area k the month m . All the parameters used are detailed in the Additional file 2.

The probability of introduction for PW-host is the probability to import at least one infected host able to transmit the infection to at least one local host and is defined as:

$$P(\text{introH}_{ijkm}) = 1 - [1 - P(\text{relH}_{ijkm}) \times P(\text{estH}_{ijkm})]^{e_{ijkm}}$$

Where $P(\text{relH}_{ijkm})$, the probability of release, depends of the importation procedure implemented and the periods where a host is infected and $P(\text{estH}_{ijkm})$, the probability of establishment, is defined as:

$$P(\text{estH}_{ijkm}) = 1 - [1 - I_{VH} \times P(\text{surv}_{km}) \times b_{equik} \times I_{HV}]^{c_{ulikm}}$$

with c_{ulikm} the number of vectors feeding on an infected viraemic imported host equals at $BR_{km} \times \text{Vir} \times C_{km}$

For each category of exporting region, there is different import procedure implemented and thus different periods z where a host can be infected. For a given region j , there is a total of w different time periods z where the equine can be infected depending on the import procedure implemented for the region j . The different periods z for each region j are presented below:

- **High risk countries:** host can be infected 1) Before quarantine, 2) During quarantine but before the first serological test CF1, 3) During quarantine but between the both serological tests CF1 and CF2, 4) During quarantine but after CF2 and before clinical exam, or 5) After clinical exam.
- **Low risk countries:**
 - o **Non EU country member:** host can be infected 1) Before quarantine, 2) During quarantine but before CF1, 3) During quarantine but between CF1 and CF2, 4) During quarantine but after CF2 and before clinical exam, or 5) After clinical exam.
 - o **EU country member:** host can be infected 1) Before clinical exam, or 2) After clinical exam.
- **Very low risk countries:**
 - o **Non EU country member:** host can be infected 1) Before clinical exam, or 2) After clinical exam.
 - o **EU country member:** host can be infected 1) Before clinical exam, or 2) After clinical exam.

The probability of release by species i from region j to area k during a specific month m ($P(\text{relH}_{ijkm})$) is thus calculated as: $P(\text{relA}_{ijkm}) = \frac{\sum_{z=1}^w (\text{length period } z) \times P(\text{relA}_{ijkzm})}{\sum_{z=1}^w (\text{length period } z)}$

Where $P(\text{relH}_{ijkzm})$ is the probability of release when the animal i is infected during the time period z . $P(\text{relA}_{ijkzm})$ is calculated for each period z as:

$$P(\text{relA}_{ijkzm}) = P(\text{inf}_{ijmz}) \times P(\text{vir}_{ijmz}) \times (1 - P(\text{CF1}_{iz})) \times (1 - P(\text{CF2}_{iz})) \times (1 - P(\text{clin}_{ijmz})) \times (1 - P(\text{trans}_{ijkz}))$$

1. $P(\text{inf}_{ijmz})$ = Probability for a host to be infected during period z in the month m in area j

The probability of infection during a certain period z (before or during the import procedure) depends on the fraction of this period z spend in each of the months m , $m-1$ and $m-2$.

a. No quarantine and CF test are required

Entire period of being at risk of infection is the high risk period (HRP).

• **Probability that the imported host is infected before clin**

If $\text{HRP} < e$

$$= \text{PO}_{jm} \times \text{CI}_m$$

If $\text{HRP} > e$

$$\text{If } \text{HRP} < 30 + e$$

$$= \frac{PO_{jm}}{HRP} \times [CI_m \times e + CI_{m-1} \times (HRP - e)]$$

If $HRP > 30 + e$

$$= \frac{PO_{jm}}{HRP} \times [CI_m \times e + CI_{m-1} \times 30 + CI_{m-2} \times (HRP - 30 - e)]$$

- **Probability that the imported host is infected after clin**

$$= PO_{jm} \times CI_m$$

b. Quarantine and CF tests required

- **Probability that the imported host is infected before q**

If $q - e < 30$

$$= \frac{PO_{jm}}{Inf_{time}} \times [CI_{m-1} \times (30 - q + e) + CI_{m-2} \times (Inf_{time} - (30 - q + e))]$$

If $q - e > 30$

$$= \frac{PO_{jm}}{Inf_{time}} \times [CI_{m-2} \times (60 - q + e) + CI_{m-3} \times (Inf_{time} - 60 + q - e)]$$

- **Probability that the imported host is infected between q and cf1**

If $e - cf1 < 0$

If $q > 30 + e$

$$= \frac{(1 - Prot_{vect}) \times PO_{jm}}{q - cf1} \times [CI_{m-1} \times (30 - cf1 + e) + CI_{m-2} \times (q - 30 - e)]$$

If $q < 30 + e$

$$= (1 - Prot_{vect}) \times PO_{jm} \times CI_{m-1}$$

if $e - cf1 > 0$

if $q > 30 + e$

$$= \frac{(1 - Prot_{vect}) \times PO_{jm}}{q - cf1} \times [CI_m \times (e - cf1) + CI_{m-1} \times 30 + CI_{m-2} \times (q - 30 - e)]$$

if $q < 30 + e$

$$= \frac{(1 - Prot_{vect}) \times PO_{jm}}{q - cf1} \times [CI_m \times (e - cf1) + CI_{m-1} \times (q - e)]$$

- **Probability that the imported host is infected between cf1 and cf2**

If $e < cf2$

$$= \frac{(1 - Prot_{vect}) \times PO_{jm}}{cf1 - cf2} \times [CI_{m-1} \times (30 - cf2 + e) + CI_{m-2} \times (cf1 - 30 - e)]$$

If $e > cf2$

if $cf1 > 30 + e$

$$= \frac{(1 - Prot_{vect}) \times PO_{jm}}{cf1 - cf2} \times [CI_m \times (e - cf2) + CI_{m-1} \times 30 + CI_{m-2} \times (cf1 - 30 - e)]$$

if $cf1 < 30 + e$

$$= \frac{(1 - Prot_{vect}) \times PO_{jm}}{cf1 - cf2} \times [CI_m \times (e - cf2) + CI_{m-1} \times (cf1 - e)]$$

- **Probability that the imported host is infected after cf2**

If $e < cf2$

$$= \frac{(1 - Prot_{vect}) \times PO_{jm}}{cf2} \times [CI_m \times e + CI_{m-1} \times (cf2 - e)]$$

If $e > cf2$

$$= (1 - Prot_{vect}) \times PO_{jm} \times CI_m$$

2. $P(vir_{ijmz})$ = Probability for a host to be vireamic or incubating when imported to area B given being infected

Calculation is based on a constant viraemic and latent period, which is equal for each equine of species i .

a. No quarantine and CF test are required

• **When infected before clin**

If $\ln + \text{Vir} < t_{AB} + \text{clin}$

= 0

If $\ln + \text{Vir} > \text{HRP} + t_{AB} + \text{clin}$

= 1

If $\ln + \text{Vir} < \text{HRP} + t_{AB} + \text{clin}$

$$= \frac{\ln + \text{Vir} - t_{AB}}{\text{HRP} - \text{clin}}$$

• **When infected after clin**

If $\ln > t_{AB} + \text{clin}$

= 1

If $\ln < t_{AB} + \text{clin}$

If $\ln + \text{Vir} > t_{AB} + \text{clin}$

= 1

If $\ln + \text{Vir} < t_{AB} + \text{clin}$

$$= \frac{\ln + \text{Vir}}{t_{AB} + \text{clin}}$$

b. Quarantine and CF tests required

• **When infected before q**

If $\ln + \text{Vir} > \text{Inf}_{\text{time}} + q + t_{AB}$

= 1

If $\ln + \text{Vir} < q + t_{AB}$

= 0

If $\ln + \text{Vir} < \text{Inf}_{\text{time}} + q + t_{AB}$

$$= \frac{\ln + \text{Vir} - q - t_{AB}}{\text{Inf}_{\text{time}}}$$

• **When infected between q and cf1**

If $\ln + \text{Vir} > q + t_{AB}$

= 1

If $\ln + \text{Vir} < \text{cf1} + t_{AB}$

= 0

If $q + t_{AB} > \ln + \text{Vir} > \text{cf1} + t_{AB}$

$$= \frac{\ln + \text{Vir} - \text{cf1} - t_{AB}}{q - \text{cf1}}$$

• **When infected between cf1 and cf2**

If $\ln + \text{Vir} > \text{cf1} + t_{AB}$

= 1

If $\ln + \text{Vir} < \text{cf2} + t_{AB}$

= 0

If $\text{cf1} + t_{AB} > \ln + \text{Vir} > \text{cf2} + t_{AB}$

$$= \frac{\ln + \text{Vir} - \text{cf2} - t_{AB}}{\text{cf1} - \text{cf2}}$$

• **When infected after cf2**

If $\ln + \text{Vir} > \text{cf2} + t_{AB}$

= 1

If $\ln + \text{Vir} < t_{AB}$

= 0

If $\text{cf2} + t_{AB} > \ln + \text{Vir} > t_{AB}$

$$= \frac{\ln + \text{Vir} - t_{AB}}{\text{cf2}}$$

3. $P(\text{clin}_{ijmz}) = \text{Probability for an infected host to be detected during importation procedure}$

a. No quarantine and CF test are required

Probability to be detected during importation procedure = Probability to be detected by clinical inspection

- **When infected before clin**

If $\text{In} > \text{HRP} - \text{clin}$

= 0

If $\text{In} < \text{HRP} - \text{clin}$

If $\text{In} + \text{Vir} < \text{HRP} - \text{clin}$

$$= \frac{\text{Vir} \times \text{Se}_{\text{clin}}}{\text{HRP} - \text{clin}}$$

If $\text{In} + \text{Vir} > \text{HRP} - \text{clin}$

$$= \frac{(\text{HRP} - \text{clin} - \text{In}) \times \text{Se}_{\text{clin}}}{\text{HRP} - \text{clin}}$$

- **When infected after clin**

= 0

b. Quarantine and CF tests required

i. $P(\text{CF1}_{iz}) = \text{Probability for an infected host to be detected by cf1}$

- **When infected before q**

If $q - \text{cf1} < \text{Sero}$

= Se

If $\text{Inf}_{\text{time}} + q - \text{cf1} < \text{Sero}$

= $1 - \text{Sp}$

If $\text{Inf}_{\text{time}} + q - \text{cf1} > \text{Sero}$

$$= \frac{(\text{Inf}_{\text{time}} - \text{Sero} + q - \text{cf1}) \times \text{Se}}{\text{Inf}_{\text{time}}} + \frac{(\text{Sero} - q + \text{cf1}) \times (1 - \text{Sp})}{\text{Inf}_{\text{time}}}$$

- **When infected between q and cf1**

If $q - \text{cf1} < \text{Sero}$

= $1 - \text{Sp}$

if $q - \text{cf1} > \text{Sero}$

$$= \frac{(q - \text{cf1} - \text{Sero}) \times \text{Se}}{q - \text{cf1}} + \frac{\text{Sero} \times (1 - \text{Sp})}{q - \text{cf1}}$$

ii. $P(\text{CF2}_{iz}) = \text{Probability for an infected host to be detected by cf2}$

Assumption: cf1 and cf2 are independent

- **When infected before q**

if $q - \text{cf2} > \text{Sero}$

= Se

If $\text{Inf}_{\text{time}} + q - \text{cf2} < \text{Sero}$

= $1 - \text{Sp}$

If $\text{Inf}_{\text{time}} + q - \text{cf2} > \text{Sero}$

$$= \frac{(\text{Inf}_{\text{time}} - \text{Sero} + q - \text{cf2}) \times \text{Se}}{\text{Inf}_{\text{time}}} + \frac{(\text{Sero} - q + \text{cf2}) \times (1 - \text{Sp})}{\text{Inf}_{\text{time}}}$$

- **When infected between q and cf1**

if $q - \text{cf2} < \text{Sero}$

= $1 - \text{Sp}$

If $q - \text{cf2} > \text{Sero}$

If $\text{cf1} - \text{cf2} > \text{Sero}$

= Se

If $\text{cf1} - \text{cf2} < \text{Sero}$

$$= \frac{(q - cf2 - \text{Sero}) \times Se}{q - cf1} + \frac{(\text{Sero} - cf1 + cf2) \times (1 - Sp)}{q - cf1}$$

- **When infected between cf1 and cf2**

If $cf1 - cf2 < \text{Sero}$

$$= 1 - Sp$$

If $cf1 - cf2 > \text{Sero}$

$$= \frac{(cf1 - cf2 - \text{Sero}) \times Se}{cf1 - cf2} + \frac{\text{Sero} \times (1 - Sp)}{cf1 - cf2}$$

iii. $P(\text{clin}) = \text{Probability for an infected host to be detected by clinical inspection}$

- **When infected before q**

If $\ln + \text{Vir} < q - \text{clin}$ or $\ln > \ln_{\text{time}} + q - \text{clin}$

$$= 0$$

If $\ln + \text{Vir} > \ln_{\text{time}} + q - \text{clin}$

If $\ln < q - \text{clin}$

$$= Se_{\text{clin}}$$

If $\ln > q - \text{clin}$

$$= \frac{(\ln - q + \text{clin}) \times Se_{\text{clin}}}{\ln_{\text{time}}}$$

If $\ln_{\text{time}} + q - \text{clin} > \ln + \text{Vir} > q - \text{clin}$

If $\ln < q - \text{clin}$

$$= \frac{(\ln + \text{Vir} - q + \text{clin}) \times Se_{\text{clin}}}{\ln_{\text{time}}}$$

If $\ln > q - \text{clin}$

$$= \frac{\text{Vir} \times Se_{\text{clin}}}{\ln_{\text{time}}}$$

- **When infected between q and cf1**

If $\ln + \text{Vir} < cf1 - \text{clin}$ or $\ln > q - \text{clin}$

$$= 0$$

If $\ln + \text{Vir} > q - \text{clin}$

If $\ln < cf1 - \text{clin}$

$$= Se_{\text{clin}}$$

If $\ln > cf1 - \text{clin}$

$$= \frac{(\ln - cf1 + \text{clin}) \times Se_{\text{clin}}}{q - cf1}$$

If $q - \text{clin} > \ln + \text{Vir} > cf1 - \text{clin}$

If $\ln < cf1 - \text{clin}$

$$= \frac{(\ln + \text{Vir} - cf1 + \text{clin}) \times Se_{\text{clin}}}{q - cf1}$$

If $\ln > cf1 - \text{clin}$

$$= \frac{\text{Vir} \times Se_{\text{clin}}}{q - cf1}$$

- **When infected between cf1 and cf2**

If $\ln + \text{Vir} < cf2 - \text{clin}$ or $\ln > cf1 - \text{clin}$

$$= 0$$

If $\ln + \text{Vir} > cf1 - \text{clin}$

If $\ln < cf2 - \text{clin}$

$$= Se_{\text{clin}}$$

If $\ln > cf2 - \text{clin}$

$$= \frac{(\ln - cf2 + \text{clin}) \times Se_{\text{clin}}}{cf1 - cf2}$$

If $cf1 - \text{clin} > \ln + \text{Vir} > cf2 - \text{clin}$

If $\ln < cf2 - \text{clin}$

$$\begin{aligned}
&= \frac{(In + Vir - cf2 + clin) \times Se_{clin}}{cf1 - cf2} \\
&\text{If } In > cf2 - clin \\
&= \frac{Vir \times Se_{clin}}{cf1 - cf2}
\end{aligned}$$

- **When infected after cf2**

If $In > cf2 - clin$

= 0

If $In < cf2 - clin$

$$\begin{aligned}
&\text{If } In + Vir > cf2 - clin \\
&= \frac{(cf2 - clin - In) \times Se_{clin}}{cf2 - clin}
\end{aligned}$$

$$\begin{aligned}
&\text{If } In + Vir < cf2 - clin \\
&= \frac{Vir \times Se_{clin}}{cf2 - clin}
\end{aligned}$$

4. $P(trans_{ijkz})$ = Probability for an infected host to be detected during transport from A to B given having passed the examinations and testing prior to embarkation.

a. No quarantine and CF test

- **When infected before clin**

If $In > HRP + t_{AB} - clin$

= 0

If $In < HRP + t_{AB} - clin$

$$\begin{aligned}
&\text{If } In + Vir < HRP - clin \\
&= \frac{Vir \times Se_{clin}}{HRP - clin} \\
&\text{If } In + Vir > HRP - clin \\
&= \frac{(HRP - clin + t_{AB} - In) \times Se_{clin}}{HRP - clin}
\end{aligned}$$

- **When infected after clin**

If $In > t_{AB} + clin$

= 0

If $In < t_{AB} + clin$

$$\begin{aligned}
&\text{If } In + Vir < clin + t_{AB} \\
&= \frac{Vir \times Se_{clin}}{t_{AB} - clin} \\
&\text{If } In + Vir > clin + t_{AB} \\
&= \frac{(clin + t_{AB} - In) \times Se_{clin}}{t_{AB} - clin}
\end{aligned}$$

b. Quarantine and CF tests required

- **When infected before q**

If $In + Vir < q + t_{AB}$ or $In > Inf_time + q + t_{AB}$

= 0

If $In + Vir > Inf_time + q + t_{AB}$

$$\begin{aligned}
&\text{If } In > q + t_{AB} \\
&= \frac{(In - q - t_{AB}) \times Se_{clin}}{Inf_time}
\end{aligned}$$

$$\begin{aligned}
&\text{If } In < q + t_{AB} \\
&= Se_{clin}
\end{aligned}$$

If $Inf_time + q + t_{AB} > In + Vir > q + t_{AB}$

$$\begin{aligned}
&\text{If } In > q + t_{AB} \\
&= \frac{Vir \times Se_{clin}}{Inf_time}
\end{aligned}$$

$$\begin{aligned} & \text{If } \ln < q + t_{AB} \\ & = \frac{(\ln - q - t_{AB}) \times Se_{clin}}{Inf_{time}} \end{aligned}$$

- **When infected between q and cf1**

$$\text{If } \ln + Vir < cf1 + t_{AB} \text{ or } \ln > q + t_{AB}$$

$$= 0$$

$$\text{If } \ln + Vir > q + t_{AB}$$

$$\begin{aligned} & \text{If } \ln > cf1 + t_{AB} \\ & = \frac{(\ln - cf1 - t_{AB}) \times Se_{clin}}{q - cf1} \end{aligned}$$

$$\text{If } \ln < cf1 + t_{AB}$$

$$= Se_{clin}$$

$$\text{If } q + t_{AB} > \ln + Vir > cf1 + t_{AB}$$

$$\text{If } \ln > cf1 + t_{AB}$$

$$= \frac{Vir \times Se_{clin}}{q - cf1}$$

$$\text{If } \ln < cf1 + t_{AB}$$

$$= \frac{(\ln + Vir - cf1 - t_{AB}) \times Se_{clin}}{q - cf1}$$

- **When infected between cf1 and cf2**

$$\text{If } \ln + Vir < cf2 + t_{AB} \text{ or } \ln > cf1 + t_{AB}$$

$$= 0$$

$$\text{If } \ln + Vir > cf1 + t_{AB}$$

$$\text{If } \ln > cf2 + t_{AB}$$

$$= \frac{(\ln - cf2 - t_{AB}) \times Se_{clin}}{cf1 - cf2}$$

$$\text{If } \ln < cf2 + t_{AB}$$

$$= Se_{clin}$$

$$\text{If } cf1 + t_{AB} > \ln + Vir > cf2 + t_{AB}$$

$$\text{If } \ln > cf2 + t_{AB}$$

$$= \frac{Vir \times Se_{clin}}{cf1 - cf2}$$

$$\text{If } \ln < cf2 + t_{AB}$$

$$= \frac{(\ln + Vir - cf2 - t_{AB}) \times Se_{clin}}{cf1 - cf2}$$

- **When infected after cf2**

$$\text{If } \ln > cf2 + t_{AB}$$

$$= 0$$

$$\text{If } \ln < cf2 + t_{AB}$$

$$\text{If } \ln + Vir > cf2 + t_{AB}$$

$$= \frac{(cf2 + t_{AB} - \ln) \times Se_{clin}}{cf2}$$

$$\text{If } \ln + Vir > cf2 + t_{AB}$$

$$= \frac{Vir \times Se_{clin}}{cf2}$$

5. $P(surv_{km})$ = Probability that the vector survives to the EIP and can have a blood meal during the month m

$$P(surv_{km}) = e^{-(N_{km} \times GC_{km} \times MR_{km})}$$

Appendix 2: Model calculation for PW-vector. Details of calculation regarding the AHSV introduction via the import of an infectious vector.

All the parameters used are detailed in the Appendix 3.

The probability to introduce a single vector from j to k during the month m which is able to induce an entire transmission cycle in which at least one local host is infected by a local vector is defined as:

$$P(\text{intro}B_{jkm}) = P(\text{rel}B_{jkm}) \times P(\text{est}B_{jkm})$$

$$\text{Where } P(\text{rel}B_{jkm}) = P(\text{trans}_{\text{culijm}}) \times P(\text{surv}_{\text{transjkm}}) \times P(\text{inf}_{\text{culijm}}) \times n_{\text{transjkm}}$$

$$\text{And } P(\text{est}B_{jkm}) = P(\text{surv}_{\text{arrivaljkm}}) \times b_{\text{equik}} \times I_{\text{VH}} \times \left[1 - \left[1 - I_{\text{VH}} \times P(\text{surv}_{\text{km}}) \times b_{\text{equik}} \times I_{\text{HV}}\right]^{\text{culikm}}\right]$$

With culikm the number of vector feeding on an infected viraemic imported host calculated as:

$$\text{culikm} = \text{BR}_{\text{km}} \times \text{Vir} \times C_{\text{km}}$$

1. $P(\text{inf}_{\text{culijm}})$ = Probability for a vector to be infected the month m in area j

$$P(\text{inf}_{\text{culijm}}) = \text{PO}_{\text{jm}} \times r_{\text{jm}}$$

2. $P(\text{trans}_{\text{culijm}})$ = Probability for a vector to be transported after infection from area j

Only a vector which is infected and transported poses a risk, therefore we only consider those vectors that are infected and transported during their life time. We assume that an infected vector will be infected at a uniformly distributed time during its life, D_{inf} . Additionally, we assume that a vector is transported at a uniformly distributed moment during its life time, which is exponentially distributed with mean $1/\text{MR}_{\text{jm}}$. The probability that the moment of transportation occurs after the infection event is equal to the part of the total lifetime of the vector that it is infected. Thus $P(\text{trans}_{\text{culijm}})$ is estimated, as made by Napp et al. (Napp et al. 2012), as:

$$P(\text{trans}_{\text{culijm}}) = \frac{(1/\text{MR}_{\text{jm}} - D_{\text{inf}})}{1/\text{MR}_{\text{jm}}} = 1 - D_{\text{inf}} \text{MR}_{\text{jm}}$$

NB: Temperature in departure area j was assumed to be constant over months and thus MR_{jm} is here also constant over months.

3. $P(\text{surv}_{\text{transjkm}})$ = Probability for a vector to stay alive from j until the arrival in area k during the month m

The conditions during travel (e.g. temperature) are assumed to not affect the viability of culicoides except when pest control is applied (worst case scenario). There is no data available on survival rate of culicoides in an unfavorable context as assumed to occur during transport. Moreover the conditions during transports have a high variability and information are impossible to collect.

The probability to stay alive until the arrival is the probability to survive until transport and during the time of transport.

$$P(\text{surv}_{\text{transjkm}}) = e^{-\text{MR}_{\text{jm}} \times (D_{\text{trans}} + t_{\text{jk}})} \times (1 - \text{Prot}_{\text{vect}})$$

4. $P(\text{surv}_{\text{arrivaljkm}})$ = The vector survives to the transport from j , the EIP and can have at least a blood meal after the end of EIP and when arrives in the area k the month m

If $T_{\text{B}} < 0$ culicoides are assumed to not survive

= 0

$$\text{If } (N_{\text{m}} \cdot \text{GC}_{\text{jm}}) > (D_{\text{transp}} + t_{\text{jk}})$$

$$= e^{\text{MR}_{\text{km}} \times (D_{\text{trans}} + t_{\text{jk}}) - N_{\text{m}} \times \text{GC}_{\text{jm}}}$$

$$\text{If } (N_{\text{m}} \cdot \text{GC}_{\text{jm}}) < (D_{\text{transp}} + t_{\text{jk}})$$

$$\text{If } t_{\text{jk}} > \text{GC}_{\text{jm}}$$

we assume that the last GC_{m} is spent half during transport and half in the arrival area k .

$$= e^{-MR_{km} \times \frac{GC_{jm}}{2}}$$

If $t_{jk} < GC_{jm}$

we assume that the last GC_m is spent half in the departure area j and half in the arrival area k .

$$= e^{-MR_{km} \times \frac{GC_{jm} - t_{jk}}{2}}$$

NB : if $T_k < T_{\min}$ (9.5°C), where T_{\min} is the minimal temperature for formulae for MR and GC (if T_k is lower, the formulae are not valid), we will use the T_{\min} in our calculus (worst case scenario).

5. $P(\text{surv}_{km})$ = Probability that the local vector survives to the EIP and can have a blood meal during the month m in the area k

$$P(\text{surv}_{km}) = e^{-(N_{km} \times GC_{km} \times MR_{km})}$$

Appendix 3: Model parameters. Description of all parameters used in the model calculation

Preliminary definitions:

- T_{km} is the average monthly temperature during the month m in the area k (details of the estimation available in the main text)
- C_{km} is the number of competent vectors feeding on one equine in area k during the month m (details of the estimation available in the main text)
- ρ_k is the ratio of the number bovines to the number of equines per area k (details of the estimation available in the main text)
- e is defined as the day of embarkation (set the 12th of each month)
- q is the length of quarantine (40 days given EU regulation)
- $clin$ is the day of the clinical exam before embarkation (equals zero, one or two days depending on the departure region given EU regulation).
- $Cf1$ is the day when the first test is performed (35 days before embarkation given EU regulation)
- $Cf2$ is the day when the second test is performed (10 days before embarkation given EU regulation)

	PARAMETER	ESTIMATION	REFERENCES
Vertebrate hosts			
I_i	Incubation period (days)	All equines: Pert(2,6,10)	(de Vos et al. 2012) (Kazeem et al. 2008)
Vir_i	Viraemic period (longer for surviving animals than for animals succumbing from disease)	Horses: Discrete({Gamma(29.75,0.20). Gamma(20.25,0.22)};{0.3,0.7}) Donkeys: Discrete({28,12};{0.9,0.1}) Zebras: Discrete({40,28};{0.99,0.01})	(de Vos et al. 2012) (Backer and Nodelijk 2011)
$Sero_i$	Time to seroconversion (days)	All equines: Uniform(10,14)	(de Vos et al. 2012)
Vectors			
GC_{km}	Length of the gonotrophic cycle during the month m (days)	$= -1.98 + 0.07217 \times T_{km} + 2516.65/T_{km}^2$	(Wittmann et al. 2002)
BR_{km}	Biting rate = reciprocal of the blood feeding interval (= Gonotrophic cycle)	$= 0.015 \times T_{km} - 0.125$	(Wittmann et al. 2002) (Backer and Nodelijk 2011)
EIP_{km}	Length of the EIP (days)	$= 0.0085 \times T_{km} - 0.0821$	(Wittmann et al. 2002)
MR_{km}	Mortality rate of the vector (days ⁻¹)	$= 0.015 \times \exp(0.063 \times T_{km})$	(de Vos et al. 2012) (Backer and Nodelijk 2011) (Wittmann et al. 2002)
N_{km}	Number of gonotrophic cycles to complete an EIP + time to next blood meal	$= \text{Roundup}(EIP_{km}/GC_{km})$	
Interaction host vector			
λ_{HV}	Probability for a vector to become infected after feeding on a viraemic host	All equines: Beta(1.05,39.6) with a mean value of 0.02	(Iacono et al. 2013) (Venter et al. 2010)
λ_{VH}	Probability for a host to become infected after being bitten by an infectious vector	All equines: Beta(6,2) with a mean value of 0.77	(de Vos et al. 2012) (Baylis et al. 2008) (Backer and Nodelijk 2011)
Export regulations			
Se	Sensitivity of the CF test	Beta(60,4)	(de Vos et al. 2012)
Sp	Specificity of the CF test	Beta(62,2)	(de Vos et al. 2012)
Se_{clin}	Sensitivity of clinical examination	Horses 0.7; Donkeys 0.1; Zebras 0.01	(de Vos et al. 2012) (Wilson et al. 2009)
$Prot_{vect}$	Efficiency of protection against vectors	Uniform (0.5,0.9)	(de Vos et al. 2012)
D_{transp}	Day of vector transportation after infection	Uniform(D_{culi_inf} ; $1/MR_{km}$)	

Departure region j			
PO_{jm}	Probability of disease occurrence	Endemic: 1 Low risk: Gamma [(15 x HRP ₂), 1/(60 x 365)] Very low risk: Gamma [HRP ₃ , 1/(61 x 365)]	(de Vos et al. 2012)
Arrival area k			
b_{equi}	Probability for a vector to bite a susceptible host	$= 1/(a \times \rho_k + 1)$ With, a the vector preference for equidae (We assume that the vector has no host preference between cattle and horse and that $a = 1$)	(Gubbins et al. 2008) (Ninio et al. 2011)(Viennet et al. 2013)

Inf_{time} = Period when a horse can be infected before the start of import procedure such as quarantine or clinical exam when there is no quarantine

If a quarantine applied: if $HRP - q > 0$, $Inf_{time} = HRP - q$ but if $HRP - q < 0$, $Inf_{time} = 0$.

If no quarantine applied: $Inf_{time} = HRP - clin$

HRP = High risk period

The HRP is the time between virus introduction and the first formal detection. In low and very low risk region, we assumed that the first infected horse won't be detected but that the secondary cases will be. Thus the time needed to detect the second case is estimated as the time required for two incubation periods plus the time till the next infectious blood meal of a vector. In low risk regions HRP₂ is assumed equal at 22 days (based on a temperature in the region j of 18°C). In very low risk region, HRP₃ is assumed equal at 60 days (based on a temperature in the region j of 12°C). As in high risk region the virus is supposed endemic, there is no real HRP₁ because an equine can be infected at any time (during or before quarantine). We thus choose to set a period of 30 days before the start of quarantine (thus 70 days before embarkation) as the earliest stage when a host can be infected.

D_{culi_inf} = Day of vector becomes infected

To estimate the day where the vector becomes infected (D_{culi_inf}), we first only take into account the *Culicoides* susceptible to the infection. When a *Culicoides* is susceptible, one blood meal on a viraemic host is assumed sufficient for this vector to become infected (Jones and Foster 1971). Assuming a uniform distribution of the viraemic host and a constant monthly temperature T_{jm} in each departure area j , the moment (or day) of *Culicoides* infection follows a Uniform distribution between 1 (the *Culicoides* is infected the first day of its life) and $1/MR_{km}$ (the *Culicoides* is infected the last day of its life).

Cl_{ijm} = Cumulative monthly number of infectious hosts i in each departure area j

Equidae have a seasonal foaling period but the foaling season depends on the geographical area considered (North or South hemisphere). We thus assumed that Cl_{ijm} was a constant for all species in all departure region j . For low and very low risk regions, Cl_{ijm} was considered as equals at 2×10^{-4} for all species based on AHSV epidemic in Spain (Rodriguez et al. 1992a; Rodriguez et al. 1992b) (de Vos et al. 2012). For high risk region, Cl_{ijm} was estimated for horses as a Pert distribution based on data from WAHID and FAO used by de Vos *et al.* (de Vos et al. 2012): Pert(4×10^{-6} , 5.02×10^{-4} , 1×10^{-3}). For donkeys and zebras, Cl_{ijm} in high risk region were respectively assumed equal at 1.2×10^{-2} and 1.6×10^{-2} based on rate of seroconversion in foals, the surviving foaling rate and the offspring rate (Barnard 1993) (Penzhorn 1985).

r_{jm} = Prevalence of infected vectors during an outbreak in the region j

We applied for the prevalence of infected vector the same process than for infected host. Thus r_{jm} is considered as a constant in all region j . In high risk region r_{jm} is assumed at 0.014 based on data from South Africa (Scheffer et al. 2012). In low risk region, as for equidae the number of infectious animal is divided by 10^{-2} between high risk regions and low risk regions, the rate was here estimated as 1.4×10^{-4} .

Appendix 4: Article in a Professional journal published in 'Bulletin épidémiologique du RESPE'. Juillet 2013.
C. Faverjon, S. Lecollinet, S. Zientara, A. Leblond. 'Peste équine, quel risque pour la France ?'
(<http://www.respe.net/node/1604>).

Peste équine, quel risque pour la France ?

mercredi 17 Juillet 2013

par Céline FAVERJON (1), Sylvie LECOLLINET (2), Stéphan ZIENTARA (2), Agnès LEBLOND (1)

Agent pathogène et historique de la maladie

Le virus de la peste équine appartient à la famille des *Reoviridae*, genre *Orbivirus*, compte 9 sérotypes différents et affecte l'ensemble des équidés. Le virus est transmis par des insectes hématophages. Les principaux vecteurs sont les *Culicoides* (notamment *Culicoides imicola*, *Culicoides bolitinos* et le groupe *Obsoletus*).

La maladie est considérée comme enzootique en Afrique Sub-saharienne où les 9 sérotypes sont présents mais de fréquentes incursions du virus ont été observées au Maghreb et dans la péninsule arabique (sérotipe 9 impliqué). La dernière introduction du virus sur le territoire européen date de 1987 en Espagne et concernait le sérotipe 4. Cet épisode a fait suite à l'importation de zèbres infectés et asymptomatiques de Namibie et a engendré la mort d'au moins 1500 chevaux. Le foyer s'est étendu au Portugal et au Maroc et a nécessité la mise en place de mesures d'isolement et de vaccinations massives qui ont coûté plus de 30 millions de dollars avant de parvenir à éradiquer la maladie.

Dernières nouvelles du monde : cas de l'Afrique du Sud

En Afrique du Sud 8 des 9 sérotypes circulent de façon enzootique. Un dispositif spécifique à ce pays avait été mis en place par les autorités vétérinaires sud-africaines en collaboration avec l'Union Européenne. Ainsi, l'UE avait demandé qu'une zone de surveillance sans vaccination autour de Cap Town et qu'une zone de protection avec vaccination autour de cette dernière ait été mise en place. Les chevaux qui devaient être exportés vers l'UE devaient subir une quarantaine (40 jours) en station confinée à Cape Town. En 2011, un foyer dans la zone de surveillance avait entraîné l'arrêt des exportations vers l'UE. En mai 2013, soit 2 ans après cet épisode, une visite de contrôle a été effectuée par les représentants de l'UE. Cependant, les conditions ne semblent pas encore réunies pour que les échanges directs de chevaux vers l'UE puissent reprendre dans des conditions satisfaisantes.

Risque d'introduction en France

Les principaux vecteurs du virus, les *Culicoides*, sont présents sur l'ensemble du territoire français. Une introduction et diffusion du virus en France est donc théoriquement possible. L'apparition en 2006 dans le nord de l'Europe du virus de la fièvre catarrhale ovine (ou Bluetongue), virus qui a un cycle épidémiologique très semblable à celui de la peste équine (même famille virale, mêmes vecteurs, même modes de transmission et caractéristiques physio-pathologiques similaires), a renforcé l'idée que le risque d'introduction et de diffusion de la peste équine en France était loin d'être nul. L'émergence en 2011 du virus Schmallenberg, lui aussi transmis par des *Culicoides*, contribue également à illustrer la potentialité d'apparition en Europe de ce type de maladies vectorielles.

Depuis l'épisode espagnol de 1987 à 1990, la réglementation européenne relative à l'importation d'équidés vivants a beaucoup évolué afin de prévenir l'introduction du virus via des hôtes infectés. Cependant, en ce qui concerne la Bluetongue, la cause de son émergence en 2006 n'a toujours pas été élucidée et les importations légales d'animaux infectés semblent être hors de cause. D'autres voies d'introduction moins aisément contrôlables sont ainsi actuellement envisagées (introduction d'un vecteur infecté via les vents ou les transports commerciaux (avions, bateaux, trains etc.), commerce illégal d'animaux...). Le virus de la peste équine pourrait emprunter des voies similaires pour parvenir jusqu'au territoire français. Il est donc important de bien garder en mémoire le tableau clinique de la maladie car

une détection précoce est un élément clef dans le contrôle de ce type d'infection dont les conséquences pourraient être dévastatrices pour l'ensemble de la filière.

Rappels : tableau clinique, diagnostic différentiel et diagnostic

La période d'incubation varie entre 2 et 14 jours et les premiers signes cliniques apparaissent généralement entre 5 et 7 jours après l'infection. La sensibilité de chaque espèce d'équidé est très variable : Chevaux, taux de létalité dépassant les 90% - Zèbres, infection typiquement asymptomatique - Anes et mules, sensibilité intermédiaire et variable selon les individus.

Quatre formes de la maladie peuvent être rencontrées :

- **Forme pulmonaire ou suraiguë (parfois foudroyante)** : hyperthermie (40-41°C), dyspnée sévère, œdème sus-orbitaire et périorbitaire, pétéchies, jetage spumeux et mort en 24 à 72 heures.
- **Forme cardiaque ou subaiguë** : hyperthermie (39-40°C), insuffisance respiratoire, péricardite exsudative, évolution en 3 à 15 jours.
- **Forme mixte** : signes semblables aux deux formes précédentes.
- **Forme atypique** : signes nerveux ou forme fébrile pure.

Diagnostic différentiel : selon les formes, encéphalites équine, premiers stades d'une piroplasmose, purpura hémorragique, artérite virale, anémie infectieuse, autres causes d'insuffisance respiratoire, d'œdème pulmonaire et de péricardite.

Diagnostic : nécropsie (type septicémique à dominante respiratoire et cardiaque, hémorragies et pétéchies viscérales), sérologie (anticorps détectables dès 10 à 15 jours après infection) ou virologie. Le seul laboratoire agréé pour effectuer ces analyses est le laboratoire de santé animale d'Alfort (notamment l'unité de virologie).

Pour aller plus loin

- Zientara S., Ponçon N., Martínez-López B., Sánchez-Vizcaíno J.M., 2012 (Avril). « La peste équine : de l'expérience espagnole au risque pour la France. ». Bulletin épidémiologique, santé animale et alimentation n°49 spécial équidés : 26 – 29. <http://www.ansespro.fr/bulletin-epidemiologique/Documents/BEP-mg-BE49-ar...>
- Mellor, Philip. Scott, et Christopher Hamblin. 2004. « African horse sickness ». Veterinary research 35 (4) (août): 445-466. doi:10.1051/vetres:2004021.
- Monographie Peste équine : http://agriculture.gouv.fr/sites/guide_epizooties/monographies/f-pe.htm

Appendix 5: Article in a Professional journal published in 'N° spécial de Pratique Vétérinaire Équine'.
Octobre 2014, 142-151. S. Zientara, C. Faverjon, A. Leblond, S. Lecollinet. La peste équine : épidémiologie,
diagnostic et prévention.

LES MALADIES INFECTIEUSES CHEZ LES ÉQUIDÉS ADULTES

—Virologie

La peste équine : épidémiologie, diagnostic et prévention

www.
lepointveterinaire.fr

Complément
de lecture : Encadré
http://www.lepoint
veterinaire.fr/
bdd/166/166_4455

Afin de lutter efficacement contre la peste équine en cas d'émergence, le praticien équin, au cœur du dispositif de surveillance, doit détecter précocement les premiers cas cliniques.



Stéphan Zientara*,
Céline Faverjon**,
Agnès Leblond***, Sylvie Lecollinet*

* UPE, UMR 1161, Anses,
Inra, ENVA, 23, avenue
du Général-de-Gaulle,
94706 Maisons-Alfort
** Inra UR346
Épidémiologie animale,
VetAgro Sup,
69280 Marcy-L'Étoile
*** Département
hippique, VetAgro Sup,
1, avenue Bourgelat,
69280 Marcy-L'Étoile

CONFLIT D'INTÉRÊTS : AUCUN

La peste équine est une maladie saisonnière des équidés, caractérisée par une évolution grave, le plus souvent mortelle, de troubles fébriles intenses, associés à des atteintes sévères des fonctions cardiaque et respiratoire. Elle est devenue une véritable entité nosologique avec l'établissement en différentes régions du continent africain d'un peuplement d'équidés européens. L'histoire de la peste équine reflète en quelque sorte l'exploration de ce continent et l'introduction concomitante d'équidés sensibles. Plusieurs épidémies spectaculaires ont porté la peste hors

de son berceau africain et attiré l'attention des autorités vétérinaires sanitaires sur la menace qu'elle constitue pour le cheptel équin mondial en raison de son pouvoir de diffusion rapide.

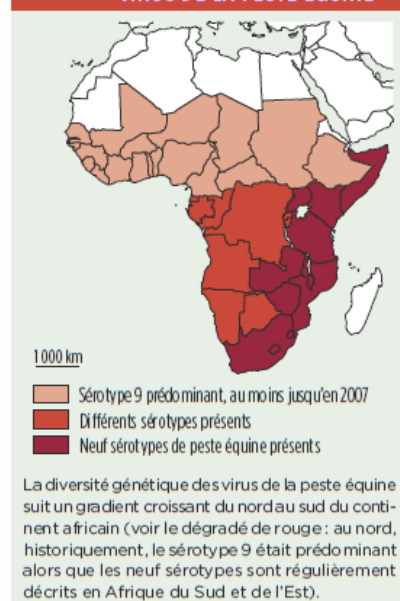
—Répartition géographique et importance économique

La peste équine est enzootique sur le continent africain au sud d'une ligne allant du Sénégal et de la Gambie à l'ouest à l'Éthiopie à l'est et jusqu'en Afrique du Sud (figure 1). La peste équine est certainement présente depuis longtemps en Afrique du Sud, mais elle n'y a été signalée qu'après la première grande épidémie de 1719, provoquant la mort de 1 700 chevaux [22]. Elle peut sortir hors de cette zone d'enzootie et provoque dans les nouvelles régions où elle apparaît des flambées épidémiques meurtrières : à titre d'exemples, les trois épidémies de 1943-1944 en Égypte, en Palestine, en Syrie, en Jordanie et au Liban, de 1959-1960 au Moyen-Orient et en Asie, qui entraîna la mort de 300 000 équidés (Chypre, Turquie, Liban, Iran, Irak, Syrie, Jordanie, Palestine, Pakistan et Inde), et de 1965-1966 dans le Maghreb et en Espagne. Ces trois épidémies ont toutes été causées par des virus de sérotype 9.

L'Europe et l'Afrique du Nord sont restées indemnes jusqu'en 1987, où un foyer, dû à un virus de sérotype 4, a été confirmé dans la province de Madrid (encadré et figure 2 complémentaires sur www.lepointveterinaire.fr) [16].

Depuis 2007, des incursions inattendues de virus de sérotypes 2 et 7 ont été décrites dans plusieurs pays d'Afrique centrale (Sénégal, Mali, Gambie, etc.) où historiquement seuls des virus de sérotype 9 avaient été isolés (figure 3). Cette expansion récente de deux sérotypes viraux aux confins des pays du Maghreb est particulièrement inquiétante,

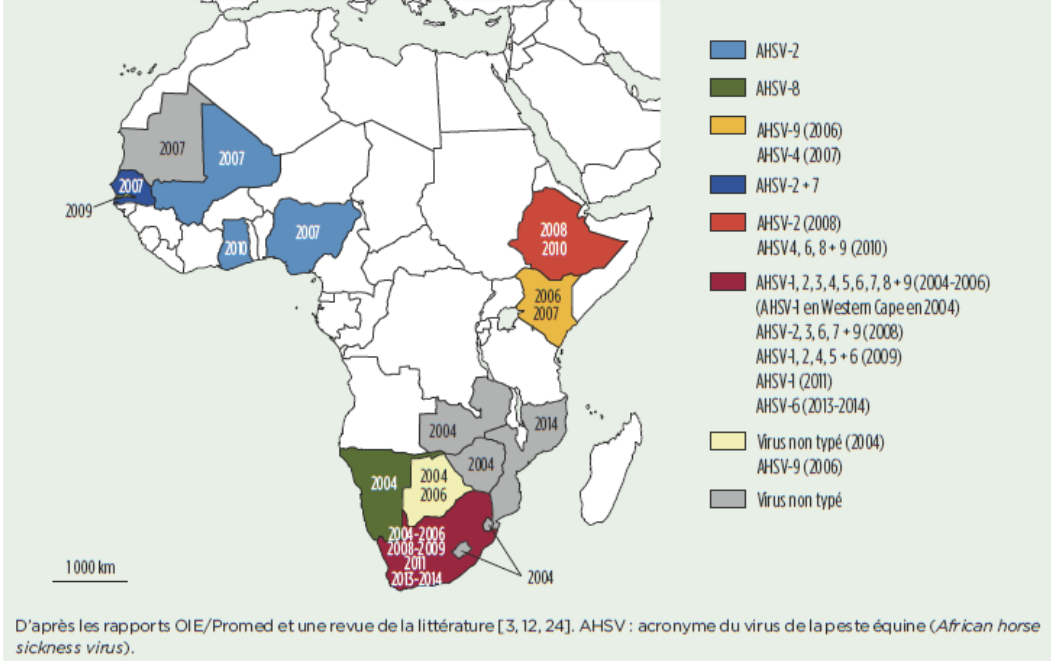
FIGURE 1 : ZONE D'ENZOOTIE POUR LE VIRUS DE LA PESTE ÉQUINE



	Éléments à retenir
--	---------------------------

- | | |
|---|---|
| <ul style="list-style-type: none"> ✦ La peste équine est une maladie virale non contagieuse présente en Afrique et transmise par des moucheron du genre <i>Culicoides</i>. ✦ La peste équine est une affection gravissime des chevaux, avec des taux de mortalité pouvant atteindre 100 % dans les formes pulmonaires de l'infection. | <ul style="list-style-type: none"> ✦ Des mesures de quarantaine et de contrôles aux frontières sécurisent les mouvements de chevaux vers l'Europe. ✦ La lutte contre des foyers de peste équine repose sur des mesures de prophylaxie sanitaire et sur une vaccination adaptée. |
|---|---|

FIGURE 3 : FOYERS RÉCENTS DE PESTE ÉQUINE (2004 À 2014)



les expériences récentes avec le virus proche de la fièvre catarrhale ovine ayant montré qu'une fois ce type de virus présent en Afrique du Nord, il peut facilement et rapidement disséminer dans le bassin méditerranéen.

—Épidémiologie

Espèces affectées

La peste équine est une maladie qui n'affecte, dans les conditions naturelles, que les équidés, et principalement le cheval, de loin le plus sensible (formes aiguës et suraiguës fréquemment mortelles), puis le mulet et le bardot (formes curables généralement), et, enfin, l'âne (formes volontiers asymptomatiques).

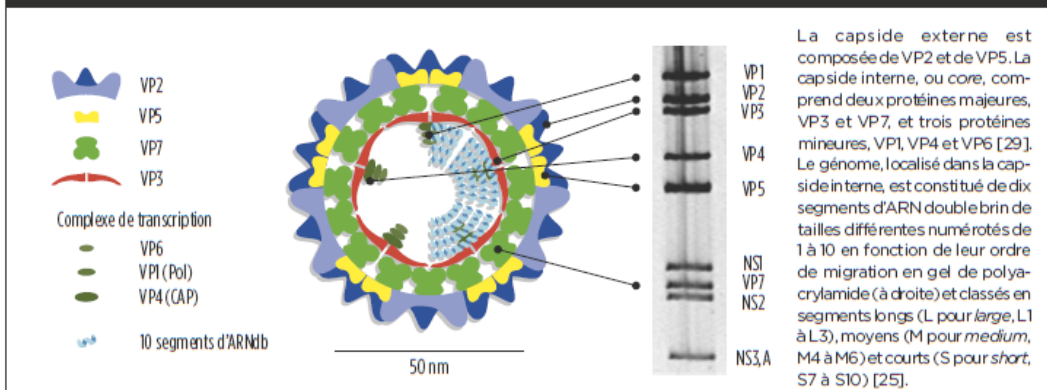
En dehors des équidés, seuls les chiens sont hautement sensibles à la peste équine et peuvent développer une forme pulmonaire fatale [21]. Ces formes restent cependant rares, de nombreux chiens présentant après infection une séroconversion sans signes cliniques. De plus, les canidés ne joueraient qu'un rôle anecdotique dans le cycle de transmission de la peste équine et sont actuellement considérés comme des culs-de-sac épidémiologiques.

Sensibilité

L'âge et le sexe ont peu ou pas d'influence sur l'évolution de l'infection : les sensibilités individuelles de chevaux âgés de 4 mois à plus de

LES MALADIES INFECTIEUSES CHEZ LES ÉQUIDÉS ADULTES

FIGURE 4 : STRUCTURE DU VIRUS DE LA PESTE ÉQUINE



15 ans, des deux sexes, sont identiques [13]. L'influence du patrimoine génétique est, en revanche, illustrée par la grande résistance des zèbres et des ânes, et par l'observation que les races de chevaux importées sont beaucoup plus réceptives et sensibles que les races locales, et ne semblent pas acquérir de résistance, même après plusieurs générations en région infectée. De plus, Laegreid et coll. ont montré que trois isolats clonés de virus équine provoquaient des formes cliniques caractéristiques (formes pulmonaire avec 100 % de mortalité en 6 à 8 jours, cardiaque avec 60 % de mortalité en 12 à 14 jours et fébrile sans mortalité) chez des chevaux qui n'ont jamais été en contact avec ce virus [14]. Ainsi, le phénotype viral est bien, expérimentalement, le facteur déterminant primaire de la forme clinique provoquée par l'infection chez ces chevaux.

Agent pathogène

Le virus équine appartient au genre *Orbivirus* de la famille des *Reoviridae*, comme le virus de la fièvre catarrhale ovine (ou BTV pour bluetongue virus).

Structure

Le virus équine possède sept protéines structurales différentes, de VP1 à VP7, réparties en deux capsides [6] (figure 4). De plus, cinq protéines non structurales, NS1, NS2, NS3 (NS3A) et NS4, sont identifiées dans les cellules infectées. Sur la base de la variabilité génétique de la protéine de capside externe VP2, neuf sérotypes du virus de la peste équine ont pu être définis par neutralisation virale [11].

Réassortiment de segments génomiques

L'organisation du génome viral en plusieurs segments génomiques autorise les échanges de

segments (ou réassortiments) entre les virus au cours d'infections simultanées d'un même animal par deux virus de la peste équine. Ce phénomène de réassortiment participe à l'évolution génétique rapide des *Orbivirus*. *In vitro*, des virus réassortis de la peste équine ont été obtenus par co-infection de cellules par des virus de sérotypes 2, 3 et 4 [18]. Des données suggèrent qu'*in vivo* les événements de réassortiment seraient plus fréquents chez l'insecte que chez l'hôte vertébré, et qu'ils sont plus difficiles à documenter chez l'hôte vertébré pour le virus de la peste équine que pour la fièvre catarrhale ovine [28].

Sources de virus, transmission

Voies de contamination

À l'exception du cas particulier des canidés généralement contaminés par ingestion de matières virulentes (viandes ou abats d'équidés infectés), la peste équine se transmet chez les équidés de façon indirecte par l'intermédiaire d'arthropodes hématophages. De nombreux vecteurs semblent potentiellement capables de transmettre la peste équine (notamment les moustiques des genres *Aedes*, *Culex* et *Anopheles*, ou les tiques des genres *Hyalomma* et *Rhipicephalus*) [20]. Cependant, le vecteur biologique majeur est un moucheron du genre *Culicoides* [8]. Les espèces de *Culicoides* (*C. imicola*, *C. bolitinos* en Afrique, *C. sonorensis* en Amérique, *C. obsoletus* et *C. pulicaris*) dont le rôle dans la transmission de la peste équine est prouvé ou très fortement suspecté sont aussi impliquées dans la transmission de la fièvre catarrhale ovine. Parmi ces espèces vectrices, *C. imicola* est décrit depuis 2000 en Corse et plus largement en Europe du Sud et *C. obsoletus* est abondant en Europe de l'Ouest et du Nord. Une quantité importante de particules virales infectieuses dans le sang de l'équidé est nécessaire

LA PESTE ÉQUINE : ÉPIDÉMIOLOGIE, DIAGNOSTIC ET PRÉVENTION

pour qu'une contamination de l'insecte vecteur puisse s'établir (titre de 10^4 à $10^{4,5}$ DMI50/0,02 ml sang pour *C. imicola*).

Réservoir de virus

L'apparition explosive de la maladie certaines années, supposant la contamination simultanée d'un grand nombre d'insectes avant l'apparition des premiers cas équins, plaide en faveur de l'existence d'un réservoir de virus présentant une virémie persistante ou à éclipse. Les chevaux ne seraient que les révélateurs d'une virose entretenue à bas bruit chez au moins un hôte réservoir (zèbres, ânes africains, etc.).

La persistance de la peste équine en Espagne pendant 4 ans pourrait, d'après Mellor, s'expliquer par la capacité de survie de *C. imicola* en hiver, saison particulièrement tempérée dans le sud de l'Espagne [19]. Une autre hypothèse avancée voudrait que les ânes ou les mules infectés aient servi de réservoirs. Ainsi le virus pourrait-il circuler entre la population d'insectes vecteurs dont la densité est excessivement réduite pendant la saison froide et la population d'ânes ou de mules chez laquelle la virémie est plus longue (virémie de 10 à 27 jours) que chez le cheval (virémie plus intense de 4 à 8 jours, au maximum 18 jours) [9, 17, 19].

Évolution

L'évolution dans le temps des épizooties de peste équine est directement liée aux périodes d'activité des vecteurs en saison chaude et humide :

- en région subtropicale, après la saison des pluies ;
- en région tempérée, dès le printemps et jusqu'à la fin de l'automne.

L'évolution dans l'espace est tributaire des zones de pullulation des vecteurs. Celles-ci se rencontrent surtout dans les régions basses et humides (marécages, bords des fleuves et cours d'eau) et à proximité des points d'eau, où le couvert végétal entretient les conditions d'hygrométrie favorables au développement des larves et à la survie des adultes.

—Pathogénie

Après inoculation par le vecteur, le virus équine se réplique initialement dans les nœuds lymphatiques régionaux. Puis il dissémine par voie sanguine, entraînant une phase de virémie intense pendant laquelle le virus est lié fortement aux hématies [5, 27].



Formes cliniques de peste équine

01. et 02. Forme respiratoire, avec un jetage séreux abondant. La présence de sang dans le jetage signe des hémorragies pulmonaires.

03. Forme cardiaque, avec des œdèmes de la face, des saillies, des paupières et de l'encolure.

Clichés : J.-P. Ganière, Oniris, et S. Zientara, Anses

Une fois dans la circulation générale, le virus se multiplie dans les cellules endothéliales et mononucléées qui constituent le site de répllication secondaire [26]. Les organes cibles sont multiples, incluant les poumons, le cœur, la rate et le tissu lymphoïde. La répllication du virus dans ces organes cibles entraîne une seconde phase de virémie, des lésions des cellules endothéliales et une activation des macrophages avec libération de cytokines (telles que l'interleukine 1 et le TNF α). Ces cytokines pro-inflammatoires et l'action directe ou indirecte du virus se traduisent par une augmentation de la perméabilité vasculaire avec transsudation du plasma dans les tissus sous-cutanés et pulmonaire et les cavités. Ce phénomène est particulièrement grave dans le parenchyme pulmonaire (œdème, jetage mousseux, asphyxie) et le cœur (défaillance cardiaque). Le tropisme des virus équine pour les cellules endothéliales pulmonaires ou cardiaques est variable, et permet d'expliquer les diverses formes cliniques observées de la maladie. Des travaux récents ont montré l'existence d'une corrélation entre le pouvoir infectieux des souches *in vitro* et l'intensité des lésions anatomo-pathologiques découvertes *in vivo* [27]. La virulence est associée à des phénomènes thrombocytopéniques, à l'augmentation du temps de coagulation et à la présence de produits de dégradation de la fibrine. Ces mécanismes sont la conséquence de la lyse et de l'activation des cellules endothéliales et des macrophages, qui aboutit à une coagulation intravasculaire disséminée.

—Symptômes

L'incubation est de durée variable selon la virulence de la souche et la réceptivité de l'équidé, en moyenne de 3 à 15 jours.

La maladie survient après une poussée fébrile irrégulière et progressivement ascendante. Elle peut ensuite évoluer sous des formes quelque peu différentes selon la prédominance de l'atteinte pulmonaire ou cardiaque (photos 1 à 3) [27].

Forme pulmonaire

C'est la forme la plus grave et la plus dramatique. Elle débute par une ascension thermique rapide

LES MALADIES INFECTIEUSES CHEZ LES ÉQUIDÉS ADULTES

(40 à 42 °C en 2 à 4 jours) associée à un syndrome fébrile avec une tachycardie, une congestion des muqueuses (parfois des pétéchies) et une anorexie plus ou moins brutale. Une sudation, diversement localisée (naseaux, base des oreilles, faces latérales de l'encolure, aine, anus, etc.), est notée chez certains individus [22]. La fréquence respiratoire s'accroît et une dyspnée s'installe : le faciès est angoissé, la langue pendante, et les naseaux sont dilatés. L'animal se tient immobile, la tête tendue sur l'encolure, les membres antérieurs écartés et le dos voûté. La tachycardie devient manifeste, le pouls très discret, signant l'évolution vers un état de choc.

La difficulté respiratoire s'accroît rapidement et un jetage séreux vient encombrer les naseaux : une toux forte, spasmodique et douloureuse secoue l'animal. Très vite, sa fréquence augmente et elle se transforme en quintes prolongées irrépressibles. Le jetage prend alors un aspect spumeux de "blanc d'œuf en neige" à cause de son brassage avec l'air dans les voies respiratoires.

À ce stade, l'animal maintient avec peine son équilibre, il se couche ou tombe brutalement et meurt par asphyxie sans agitation. Dans les minutes qui précèdent la mort, de grandes quantités de jetage mousseux peuvent s'écouler des naseaux. Chez certains individus, il s'écoule moins d'une demi-heure entre l'apparition de la dyspnée et la mort. Cependant, en général, l'issue fatale se produit en 24 à 48 heures.

Forme œdémateuse ou cardiaque

Elle se rencontre chez les individus plus résistants ou infectés par une souche de pouvoir pathogène plus faible [4].

La poussée thermique initiale est, ici, plus progressive et moins intense (acmé à 39 à 40 °C atteint en 10 à 12 jours). Elle se maintient ensuite en plateau ou, le plus souvent, diminue progressivement. Les signes généraux sont plus discrets et l'appétit peut être conservé.

Vers le 14^e à 15^e jour, alors que la baisse de température est amorcée, apparaissent des œdèmes sous-cutanés. Ils débutent dans les fosses temporales par une déformation en saillie de la région sus-orbitale, qui peut atteindre le volume d'une mandarine en 3 à 4 jours. La précocité d'apparition de ces œdèmes en cours de phase fébrile est un élément de gravité du pronostic.

Parfois, ce gonflement disparaît en quelques jours. Il peut cependant persister, s'étendre et

atteindre les paupières qui, tuméfiées, se ferment avec parfois une éversion de la conjonctive. Le globe oculaire est exorbité par l'œdème sous-jacent, il fait saillie et devient fixe. Un épiphora abondant souille les joues, les régions des masséters et intermandibulaire, le chanfrein et les naseaux. L'œdème atteint parfois le larynx et un bruit de cornage, principalement inspiratoire au début, peut être audible. La tête présente alors un aspect tuméfié. Dans certains cas, l'œdème envahit aussi l'encolure et descend le long des membres antérieurs. Il s'agit d'un œdème froid, indolore, ferme au début : le "signe du godet" n'apparaît qu'en quelques jours.

Les bruits du cœur deviennent plus faibles en raison de la formation d'une péricardite exsudative. Le pouls, jusque-là fort et bien frappé, devient filiforme et imperceptible.

Dans les cas graves, le processus œdémateux peut s'étendre à l'appareil respiratoire (accélération du rythme, expiration biphasique et phase expiratoire augmentée), mais la dyspnée n'est jamais aussi intense que dans la forme pulmonaire, et la toux et le jetage demeurent absents. L'animal, jusque-là apathique, finit par se coucher. L'état de choc s'aggrave, et l'apparition de sueurs froides, le refroidissement des oreilles, des mouvements désordonnés (simulant des coliques) et une détresse respiratoire annoncent l'arrêt cardiaque.

L'évolution mortelle se fait en 3 à 10 jours après le développement des œdèmes sous-cutanés, mais la guérison est également possible quelle que soit l'importance des œdèmes sous-cutanés.

Forme intermédiaire

Dans ce cas, les signes pulmonaires et les œdèmes sous-cutanés apparaissent simultanément ou successivement dans un ordre indéterminé. La défaillance cardiaque ou l'asphyxie emporte l'animal. L'examen nécropsique révèle la coexistence très fréquente des atteintes pulmonaire et cardiaque malgré un tableau clinique plus sélectif.

Formes fébriles

Une réaction thermique (hyperthermie de 39 à 40 °C) qui s'accompagne d'une légère polypnée, d'une tachycardie, d'une inappétence et d'une nonchalance s'estompe en 10 à 15 jours avec retour à un habitus normal.

—Examens complémentaires

Examens de laboratoire

Quelle que soit la forme clinique, les résultats des numération et formule sanguines montrent une leucopénie caractérisée par une neutropénie, une thrombocytopénie et une hémococoncentration. La biochimie sanguine révèle des anomalies non spécifiques, avec, selon les cas, une élévation de la créatine kinase, de la lactate déshydrogénase, de la phosphatase alcaline, de la créatinine et/ou de la bilirubine.

Radiographies, échographie

Les radiographies des poumons permettent d'objectiver l'œdème pulmonaire, et l'échographie thoracique de quantifier les épanchements pleuraux et/ou péricardique.

—Lésions

Lésions macroscopiques

Le tableau nécropsique est de type septicémique à dominante œdémateuse à la fois respiratoire et cardiaque [4].

Forme pulmonaire

Les lésions essentielles intéressent la cavité thoracique qui, à l'ouverture, apparaît totalement remplie par des poumons turgescents. La plèvre viscérale est luisante, humide, épaissie, parfois semée de pétéchies, et présente des plaques gélatineuses ou fibrineuses, surtout près de la base du cœur et autour des vaisseaux du hile. Le parenchyme pulmonaire est ferme, très humide, d'aspect irrégulier et bosselé en raison de la saillie des cloisons interlobulaires gorgées de sérosité. Des foyers emphysémateux déforment son bord ventral. Une sérosité claire, rose pâle, s'écoule abondamment à la coupe et un liquide blanc mousseux s'échappe à la pression. Les bronches, la trachée, le larynx et les cavités nasales sont encombrés d'une spumosité blanchâtre recouvrant une muqueuse congestionnée porteuse de pétéchies. Les nœuds lymphatiques sont hypertrophiés, infiltrés, enrobés d'un œdème gélatineux qui s'étale dans le médiastin, porteur de suffusions ou d'ecchymoses. La cavité pleurale renferme un épanchement clair, jaunâtre, plus ou moins abondant (de quelques millilitres à plusieurs litres). La muqueuse de l'estomac, vers le cul-de-sac glandulaire et la région pylorique, est épaissie

par l'œdème, congestionnée (de façon diffuse ou par plaques) et des lésions hémorragiques (notamment en région fundique) sont observées. Le foie, la rate et les reins sont congestionnés, tuméfiés à des degrés divers (photo 4).

Forme cardiaque

Les lésions essentielles intéressent le tissu conjonctif sous-cutané et l'appareil cardiovasculaire. Dans les tuméfactions, les tissus conjonctif et conjonctivo-adipeux sont imprégnés d'une sérosité gélatineuse pouvant infiltrer les différents tissus ou espaces de la tête, de l'encolure et de la région axillaire. La section des muscles de la tête et du cou laisse exsuder un liquide jaunâtre, même en région profonde. De plus, une péricardite exsudative bien développée est de règle. La graisse épicaudique d'aspect hémorragique peut être remplacée par un œdème gélatineux. L'épicarde et l'endocarde sont le siège d'hémorragies diffuses ou localisées. Ces dernières se retrouvent également au sein du myocarde qui présente des œdèmes et une myosite dégénérative avec des zones de nécrose focale.

Les lésions digestives, hépatiques, spléniques et rénales sont identiques à celles de la forme pulmonaire, mais souvent plus accusées.

Forme mixte

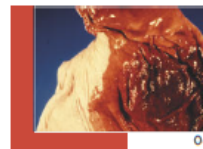
Cette forme est caractérisée par la coexistence des lésions pulmonaires, cardiaques, œdémateuses et digestives précédemment décrites.

Lésions microscopiques

Il n'existe aucune lésion microscopique caractéristique de la peste équine. Des lésions histologiques de congestion, d'œdème et d'hémorragie en rapport avec les lésions macroscopiques sont observées.

—Recommandations thérapeutiques et pronostic

Il n'existe pas de traitement antiviral spécifique. Le traitement de soutien induit l'immobilisation en box et la favorisation de la diurèse pour contrôler l'œdème pulmonaire. Cependant, il ne semble pas influencer sur l'évolution d'aucune des formes cliniques, et compte tenu de la gravité de la maladie et du risque d'infection de nouveaux mouchers à partir des équidés infectés ou malades, le traitement de la peste équine n'est pas recommandé. La morbidité et la mortalité associées à l'infection varient selon l'espèce et le statut immunitaire des



04. Hémorragies dans la région fundique de l'estomac.
Cliché : S. Zientara, Anses

LES MALADIES INFECTIEUSES CHEZ LES ÉQUIDÉS ADULTES

animaux infectés. Les chevaux sont les plus sensibles, avec un taux de mortalité de 50 à 95 %, selon la forme clinique présentée. La forme pulmonaire est invariablement fatale, le taux de mortalité est supérieur à 80 % pour la forme mixte, de 50 à 70 % pour la forme cardiaque, et les chevaux qui présentent seulement de la fièvre meurent rarement. De plus, le taux de mortalité est d'environ 50 % chez les mules et de 5 à 10 % chez les ânes européens et asiatiques. Les ânes africains et les zèbres meurent rarement des suites de cette infection.

— Diagnostic

Critères de suspicion

En plus de la fièvre élevée, les éléments de l'anamnèse, et en particulier un historique récent de voyage dans un pays à risque pour le cheval malade ou l'un de ses congénères dans l'écurie ou une écurie voisine, sont déterminants pour établir la suspicion. Les autres facteurs de risque incluent la proximité avec un aéroport ou un port, associée à la présence de vecteurs potentiels dans la région.

Diagnostic clinique et anatomo-pathologique

Les symptômes et les lésions de la peste équine ne sont pas pathognomoniques, mais les tableaux clinique et nécropsique, ainsi que le caractère épidémiologique de la maladie permettent d'orienter le diagnostic.

Diagnostic différentiel

La peste équine peut cependant être confondue, notamment au début de la maladie, avec certaines formes d'anémie infectieuse équine ou d'artérite virale, l'anaplasmose granulocytaire, la pneumonie équine à virus Hendra (maladie exotique), l'infection par le virus de l'encéphalose équine (maladie exotique), la babésiose ou la theilériose, la fièvre charbonneuse, le purpura hémorragique ou encore une défaillance cardiaque congestive. Face à une infection d'allure épidémique, la piste toxicologique est également investiguée, avec en particulier la recherche dans l'environnement ou dans l'alimentation du cheval de substances toxiques entraînant des défaillances cardiovasculaires ou respiratoires, ou des morts subites. L'intoxication par l'if, la contamination de l'aliment par des ionophores ou encore l'ingestion d'hypoglycine A, présente dans les graines de plusieurs arbres du genre *Acer* (c'est-à-dire

les érables), pourront être évoquées dans cette hypothèse.

Diagnostic de laboratoire

Le recours au laboratoire est donc nécessaire afin de confirmer le diagnostic et identifier le sérotype du virus (crucial pour la mise en place des mesures de prophylaxie médicale).

Prélèvements

Le virus peut être isolé à partir d'un échantillon de tissu splénique, de poumon, de cœur ou, chez l'animal virémique, de sang total prélevé sur anticoagulant (EDTA) et conservé à + 4 °C.

Mise en évidence de l'agent pathogène

♦ L'isolement du virus doit être tenté le plus tôt possible. Si la virémie atteint un titre maximal pendant la période fébrile initiale, les chances d'isolement diminuent dès l'apparition et l'évolution des signes cliniques.

L'isolement du virus est classiquement effectué :
- soit par inoculation à des cultures de cellules. Les suspensions d'hématies lavées sont inoculées à des cultures de cellules VERO ou BHK21. L'effet cytopathique apparaît 3 à 7 jours après inoculation ou après plusieurs passages en aveugle. L'isolement et l'identification du virus nécessitent un délai minimal de 7 à 14 jours ;

- soit par inoculation intracérébrale à des souris nouveau-nées. Après une période d'incubation de 4 à 20 jours, les souris présentent des signes de prostration, de parésie, d'incoordination motrice, et meurent en 4 à 5 jours [11]. L'inoculation aux œufs embryonnés peut aussi être employée.

♦ L'identification du virus peut faire appel à des réactions sérologiques utilisant le virus isolé comme antigène :

- la fixation du complément ou l'immunofluorescence directe ou indirecte permettent l'identification du virus en culture de tissus. Ce test est spécifique de groupe. Il peut être effectué à partir de tissu splénique, mais manque de sensibilité ;
- le test de neutralisation virale à l'aide d'antisérums spécifiques de type permet l'identification du type de virus en cause. Il n'y a pas de neutralisation croisée entre les neuf virus, à l'exception des sérotypes 6 et 9 (mais aussi, dans une moindre mesure, entre les sérotypes 1-2, 3-7 et 5-8) ;
- des techniques récentes de *polymerase chain reaction* en temps réel (RT-PCR) permettent

désormais d'identifier et/ou de typer le virus en quelques heures. Elles amplifient soit des segments génomiques conservés (1, 3, 7) pour mettre en évidence un virus équinestique, soit des segments hautement variables (segment 2, VP2) pour le typage [1, 3, 24].

Diagnostic sérologique

Le diagnostic biologique d'une suspicion de peste équine ne doit pas être fondé sur l'utilisation des tests sérologiques, mais sur l'isolement du virus. En effet, de nombreux chevaux vont mourir avant d'avoir développé une réponse immunitaire à médiation humorale puisque les anticorps n'apparaissent que 10 à 14 jours après l'infection. Cependant, les techniques sérologiques permettent, à l'échelon d'un pays ou d'une région et en l'absence d'isolement du virus, de confirmer la suspicion d'une épidémie.

♦ La réaction de fixation du complément est une des méthodes sérologiques de référence recommandées par l'Organisation mondiale de la santé animale (OIE) [23]. La présence d'anticorps fixant le complément dans le sérum d'un animal traduit l'évolution d'une infection récente par un virus appartenant au groupe de la peste équine et peut aider à la confirmer du vivant de l'animal lorsque les techniques de diagnostic direct sont peu performantes. Cependant, l'anticomplémentarité des sérums d'ânes et de certains sérums de chevaux rend l'utilisation de cette méthode parfois délicate, voire impossible.

♦ Les anticorps neutralisants sont décelés à partir de la troisième semaine et persistent pendant plusieurs années.

♦ Des tests Elisa indirects ou par compétition, qui sont largement utilisés dans les laboratoires de diagnostic, ont été développés ces 20 dernières années. Ils mettent en œuvre des anticorps polyclonaux ou monoclonaux et des antigènes constitués de virus purifié ou de protéines exprimées en système procaryote ou eucaryote. Ces techniques détectent des anticorps de groupe et permettent d'identifier spécifiquement une infection ou une vaccination par un virus équinestique [21]. Laviada et coll. ont développé un test Elisa qui détecte les anticorps anti-NS3 [15]. Ces anticorps permettraient de différencier les animaux infectés ou vaccinés avec un vaccin vivant atténué (présence d'anticorps) des animaux vaccinés avec un vaccin inactivé (absence d'anticorps anti-NS3).

—Prophylaxie

La lutte contre la peste équine consiste à empêcher l'introduction de la maladie dans une région ou un pays (prophylaxie sanitaire), ou à limiter son extension à partir d'un foyer déclaré (prophylaxie médico-sanitaire).

Prophylaxie sanitaire

Elle comprend un ensemble de mesures qui ont pour objectifs de prévenir l'introduction du virus dans un pays indemne, d'empêcher l'extension d'une épidémie à des régions voisines indemnes, de limiter, de circonscrire et d'isoler les foyers de la maladie et d'en assurer l'éradication.

En région accidentellement infectée, l'abattage des animaux atteints et contaminés, la destruction rationnelle de leurs cadavres et une désinsectisation sont mis en œuvre dans les foyers. Le *Code zoosanitaire international* de l'OIE définit les notions de pays ou de région indemne de peste équine ou infecté par la maladie, ainsi que les garanties sanitaires exigées lors de l'importation d'équidés en provenance de ces zones. Des directives de l'Union européenne précisent les conditions de police sanitaire régissant les mouvements et les importations d'équidés en provenance des pays tiers, les règles de contrôle et les mesures de lutte contre la peste équine. L'autorisation d'importation en provenance de pays infectés est conditionnée à la mise en œuvre d'une quarantaine et à l'obtention de deux résultats négatifs à deux épreuves sérologiques (à un intervalle d'au moins 21 jours et d'au plus 30 jours), la seconde épreuve étant réalisée 14 jours au plus tard avant le chargement des animaux.

Prophylaxie médicale

Elle consiste en la protection des espèces sensibles par une immunisation spécifique active (vaccins monovalents ou polyvalents) ou passive (sérophylaxie). Il convient de déterminer le sérotype en cause avant toute immunisation, puisqu'il n'existe pas ou peu de protection croisée entre les neuf sérotypes du virus de la peste équine. La prophylaxie médicale repose, à l'heure actuelle, sur l'emploi de vaccins vivants atténués obtenus par passage des souches vaccinales en culture cellulaire. Ces vaccins cellulaires sont encore largement utilisés, à l'image des vaccins atténués multivalents produits par OBP (Onderstepoort

LES MALADIES INFECTIEUSES CHEZ LES ÉQUIDÉS ADULTES

Biologicals Products, Afrique du Sud) et permettant de vacciner contre huit des neuf sérotypes décrits (1 à 4 et 6 à 9). Cependant, ils présentent quelques inconvénients :

- la réversion éventuelle de la souche vaccinale à la virulence (chez l'hôte ou l'insecte vecteur) [5] ;
- le développement d'une virémie postvaccinale suffisante pour permettre l'infection du vecteur ;
- le risque potentiel d'apparition de virus réassortis par recombinaison lors de la multiplication simultanée chez un même animal de plusieurs virus vaccinaux ou de virus vaccinal et de virus sauvage ;
- le faible pouvoir antigénique de certains sérotypes (sérotipe 4 notamment) utilisés dans ces vaccins, nécessitant des immunisations répétées ;
- l'impossibilité de différencier les anticorps post-infectieux et postvaccinaux ;
- des effets tératogènes possibles, interdisant leur utilisation chez la poulinière.

En conséquence, les vaccins vivants atténués peuvent être utilisés dans l'urgence, pour contrôler une infection émergente en Europe, mais ne présentent pas une sécurité satisfaisante pour les employer dans un objectif d'éradication de la maladie.

Lors de l'épizootie équine à la fin des années 1980 en Espagne et au Portugal, un vaccin à virus inactivé monovalent a été développé par inactivation de la souche vaccinale de sérotipe 4 d'Afrique du Sud [7]. Ce vaccin a induit une protection satisfaisante et de longue durée, mais n'est plus disponible actuellement.

De plus, des candidats vaccins en cours de développement semblent prometteurs. Ils sont fondés sur l'administration d'un vecteur pox-viral exprimant les protéines VP2, seules ou en association avec VP5, ou de pseudo-particules virales constituées après autoassemblage des protéines de capside VP2, VP5 et VP7 exprimées en système baculovirus [2, 10, 25]. Les antigénicité et innocuité de ces deux types de vaccins apparaissent très satisfaisantes.

Prophylaxie médico-sanitaire

L'inefficacité des seules mesures sanitaires impose l'association des dispositifs des prophylaxies sanitaire et médicale en pays infecté. Les mesures de prophylaxie sanitaire sont appliquées dans les foyers de la maladie (abattage des animaux malades, désinsectisation, interdiction des mouvements). Autour des foyers, la vaccination des équidés est obligatoire. Ces mesures de prophylaxie médico-sanitaire

ont été appliquées en Espagne, au Portugal et au Maroc, et ont finalement permis l'éradication de l'affection au début des années 1990.

En France, la peste équine est considérée comme un danger sanitaire de première catégorie et entraîne le déploiement par les autorités sanitaires d'un plan national d'intervention sanitaire d'urgence. Dès que l'existence de foyer(s) de peste équine est confirmée, le préfet prend un arrêté portant déclaration d'infection (APDI) qui prévoit la mise en œuvre des mesures suivantes :

- dans l'exploitation, l'euthanasie sans délai des équidés malades et la destruction de leurs cadavres (l'euthanasie peut être étendue, selon les circonstances épidémiologiques, à tous les animaux présents dans le foyer) ;
- des mesures de restriction des mouvements

à l'ensemble des exploitations situées dans un rayon de 20 km autour du troupeau infecté, ainsi qu'aux exploitations en lien épidémiologique ;

- la vaccination systématique de tous les équidés se trouvant dans cette zone de 20 km.

De plus, la réglementation française prévoit la définition d'une zone de protection d'un rayon d'au moins 100 km autour de l'exploitation infectée et d'une zone de surveillance distante d'au moins 50 km du périmètre de la zone de protection. Dans ces zones, toutes les exploitations détenant des équidés sont recensées et périodiquement visitées. La vaccination des équidés peut être rendue obligatoire dans tout ou partie de la zone de protection, mais elle est interdite dans la zone de surveillance.

Conclusion

L'incursion du virus de la peste équine en Espagne à la fin des années 1980 a servi de leçon et la réglementation européenne relative à l'importation d'équidés vivants a été renforcée. Ces adaptations de la réglementation ont jusque-là porté leurs fruits puisque aucun nouveau cas n'a depuis été rapporté sur le sol européen. Cependant, l'introduction, en 2006 dans le nord de l'Europe, du BTV, malgré les mesures de protection existantes, a ravivé les inquiétudes et remis au goût du jour la question d'une possible réapparition de la peste équine sur le Vieux Continent. Ces craintes sont d'autant plus importantes qu'à ce jour le mode d'introduction du BTV n'a toujours pas été clairement identifié (les importations légales d'animaux vivants semblent être hors de cause et d'autres voies d'introduction moins aisément

contrôlables sont actuellement envisagées, comme l'introduction d'un vecteur infecté *via* les vents ou les transports commerciaux, le commerce illégal d'animaux, etc.). Cette épidémie de BTV illustre bien les possibilités d'introduction et de diffusion de maladies transmises par des *Culicoides* dans

l'espace européen. Ainsi, il convient de rester vigilant face à cette menace dont les conséquences pourraient être désastreuses. Cela est notamment vrai pour la France qui organise en 2014 les Jeux équestres mondiaux au mois d'août, en pleine période d'activité vectorielle. //

RÉSUMÉ/SUMMARY

La peste équine constitue un risque sanitaire majeur pour l'élevage du cheval en zone d'endémie (Afrique subsaharienne) et a été à l'origine à plusieurs reprises de grandes épidémies en région méditerranéenne (Afrique du Nord et Europe du Sud, notamment), à la faveur d'échanges d'équidés. L'émergence inattendue d'un virus proche, celui de la fièvre catarrhale ovine, dans le nord de l'Europe, en 2006, rend légitime la crainte d'une introduction du virus de la peste équine sur le territoire français. Afin de lutter efficacement contre la peste équine en cas d'émergence, le praticien équin, au cœur du dispositif de surveillance, devra détecter précocement les premiers cas cliniques.

Mots clés : peste équine, *Culicoides*, Afrique.

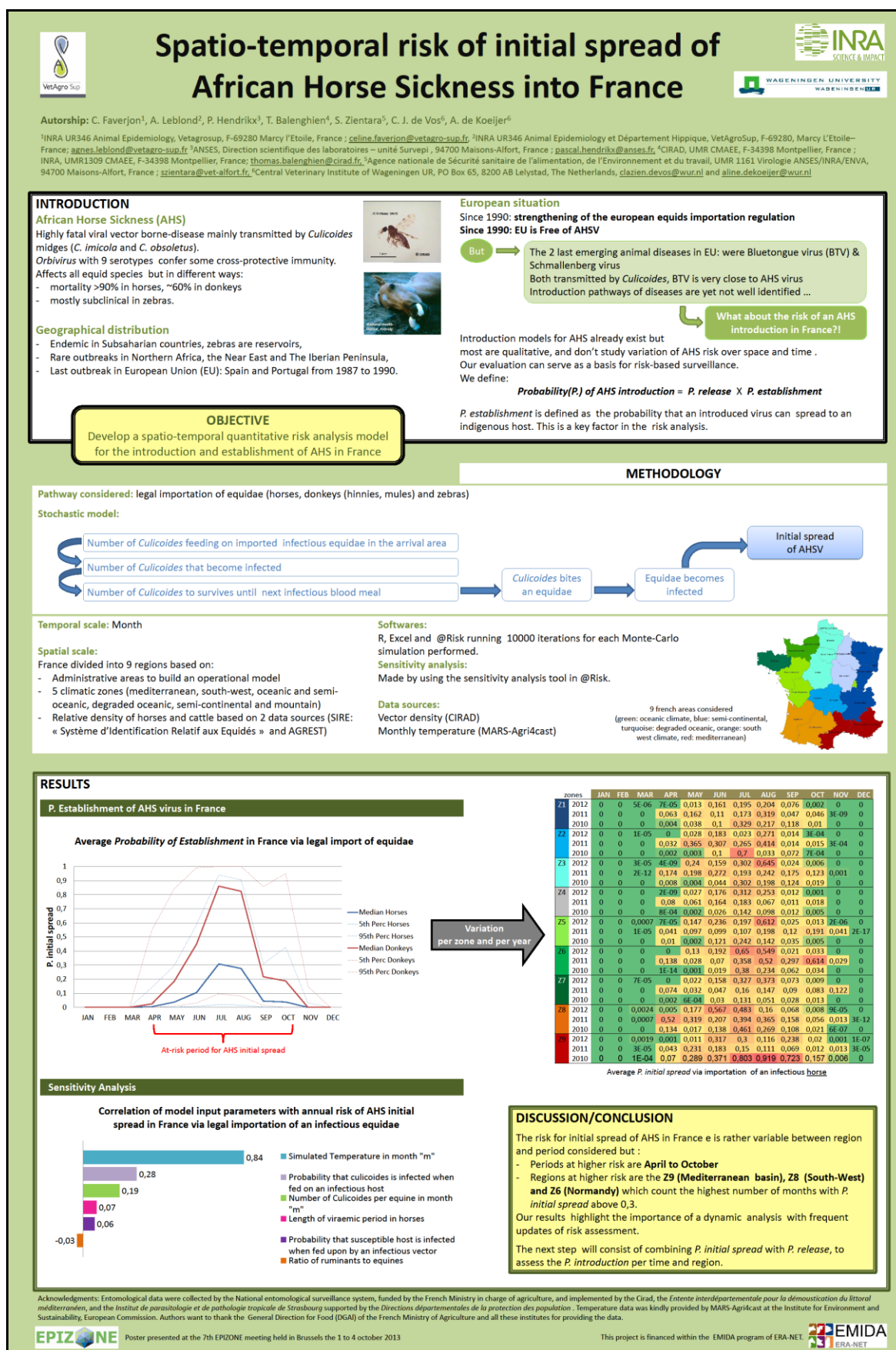
AFRICAN HORSE SICKNESS: EPIDEMIOLOGY, DIAGNOSIS AND PREVENTION

African horse sickness (AHS) is a devastating disease of equids caused by an arthropod-borne virus belonging to the Reoviridae family, genus Orbivirus. AHS is considered as a major sanitary threat for horses in endemic areas in sub-Saharan Africa. AHS virus repeatedly caused large epizootics in the Mediterranean region (North Africa and Southern Europe in particular), thanks to the trade of infected equids. The unexpected emergence of a closely related virus, the bluetongue virus, in northern Europe in 2006 has raised the fears of AHS virus introduction into Europe. To effectively control future AHS incursions, equine practitioners, key actors in the surveillance system, will have to early detect the first clinical cases.

Keywords: African horse sickness, *Culicoides* midges, Africa.

- Agüero M, Gómez-Tejedor C, Angeles Cubillo M et coll. Real-time fluorogenic reverse transcription polymerase chain reaction assay for detection of African horse sickness virus. *J. Vet. Diagn. Invest.* 2008;20(3):325-328.
- Alberca B, Bachanek-Bankowska K, Cabana M et coll. Vaccination of horses with a recombinant modified vaccinia Ankara virus (MVA) expressing African horse sickness (AHS) virus major capsid protein VP2 provides complete clinical protection against challenge. *Vaccine*. 2014;S0264-410X(14)00562-3.
- Bachanek-Bankowska K, Maan S, Castillo-Olivares J et coll. Real time rt-PCR assays for detection and typing of African horse sickness virus. *PLoS One*. 2014;9(4):e93758.
- Brown CC, Mebus CA. The pathology of African horse sickness. In: *Bluetongue, African Horse sickness and related Orbiviruses*. Eds. Walton TE, Osburn BI. CRC Press, Boca Raton, Florida. 1992:832-834.
- Burroughs TG, Laegreid WW. African Horse sickness: pathogenesis and immunity. *Comp. Immunol. Microbiol. Infect. Dis.* 1994;17(3-4):275-285.
- Burroughs JN, O'Hara RS, Smale CJ et coll. Purification and properties of virus particles, infectious sub-viral particles, cores and VP7 crystals of African horse sickness virus serotype 9. *J. Gen. Virol.* 1994;75:1849-1857.
- Dubourget P, Préaud JM, Detraz F et coll. Development, production and quality control of an industrial inactivated vaccine against African Horse sickness virus serotype 4. In: *Bluetongue, African Horse sickness and related Orbiviruses*. Eds. Walton TE, Osburn BI. CRC Press, Boca Raton, Florida. 1992:874-885.
- Du Toit RM. The transmission of bluetongue and horse sickness by *Culicoides*. *Onderstepoort J. Vet. Sci. Anim. Ind.* 1944;19:7-16.
- El Hasnaoui H, El Harrak M, Zientara S et coll. Serological and virological responses in mules and donkeys following inoculation with African Horse sickness virus serotype 4. *Arch. Virol.* 1998;142:29-36.
- Guthrie AJ, Quan M, Lourens CW et coll. Protective immunization of horses with a recombinant canarypox virus vectored vaccine co-expressing genes encoding the outer capsid proteins of African horse sickness virus. *Vaccine*. 2009;27(33):4434-4438.
- Howell PG. The isolation and identification of further antigenic types of African horse sickness virus. *Onderstepoort J. Vet. Res.* 1962;29:139-149.
- Koekemoer JJ. Serotype-specific detection of African horse sickness virus by real-time PCR and the influence of genetic variations. *J. Virol. Methods*. 2008;154(1-2):104-110.
- Laegreid WW. African Horse sickness. In: *Virus infection of equines*. Ed. Studdert MJ. Virus infection of Vertebrates. Elsevier Science Publisher. 1997:101-123.
- Laegreid WW, Skowronek A, Stone-Marschat M et coll. Characterization of virulence variants of African horse sickness virus. *Virology*. 1993;195(2):836-839.
- Laviada MD, Arias M, Sanchez-Vizcaino JM. Characterization of African horse sickness virus serotype 4-induced polypeptides in Vero cells and their reactivity in Western immunoblotting. *J. Gen. Virol.* 1993;74:81-87.
- Lubroth J. African Horse sickness and the epizootic in Spain 1987. *Equine Pract.* 1988;10:26-33.
- Lubroth J. The complete epidemiologic cycle of African Horse sickness: our incomplete knowledge. In: *Bluetongue, African Horse sickness virus and related Orbiviruses*. Eds. Walton TE, Osburn BI. CRC Press, Boca Raton, Florida. 1992:197-204.
- Meiring TL, Huismans H, van Staden V. Genome segment reassortment identifies non-structural protein NS3 as a key protein in African horse sickness virus release and alteration of membrane permeability. *Arch. Virol.* 2009;154(2):263-271.
- Mellor PS. African Horse sickness: transmission and epidemiology. *Vet. Res.* 1993;24(2):199-212.
- Mellor PS. Epizootiology and vectors of African Horse sickness virus. *Comp. Immunol. Microbiol. Infect. Dis.* 1994;17(3-4):287-296.
- Mellor PS, Hamblin C. African Horse sickness. *Vet. Res.* 2004;35(4):445-466.
- Mornet P, Gilbert Y. La peste équine. L'Expansion éditeur, coll. Les maladies animales à virus, n° 476. 1968:195p.
- OIE. Manual of standards for diagnostic tests and vaccines. 4th ed. Office international des épizooties, Paris. 2000.
- Quan M, van Vuuren M, Howell PG. Molecular epidemiology of the African horse sickness virus S10 gene. *J. Gen. Virol.* 2008;89:1159-1168.
- Roy P, Mertens PP, Casal I. African Horse sickness virus structure. *Comp. Immun. Microbiol. Infect. Dis.* 1994(3-4):17:243-273.
- Skowronek AJ, La Franco L, Stone-Marschat MA et coll. Clinical pathology and hemostatic abnormalities in experimental African horse sickness. *Vet. Pathol.* 1995;32(2):112-121.
- Stern AW. African Horse sickness. *Compend. Contin. Educ. Vet.* 2011;33(8):E1-5.
- von Teichman BF, Smit TK. Evaluation of the pathogenicity of African horse sickness (AHS) isolates in vaccinated animals. *Vaccine*. 2008;26(39):5014-5021.
- Wilson A, Mellor PS, Szmargad C et coll. Adaptive strategies of African Horse sickness virus to facilitate vector transmission. *Vet. Res.* 2009;40(2):16.

Appendix 6: Poster presented at the 7th EPIZONE meeting in Brussel, 2013.



Appendix 7: Poster presentation at the Conference of the Society for Veterinary Epidemiology and Preventive Medicine (SVEPM) in Ghent, 2015.

Quantitative RA of equine encephalosis virus (EEV) introduction into the Netherlands

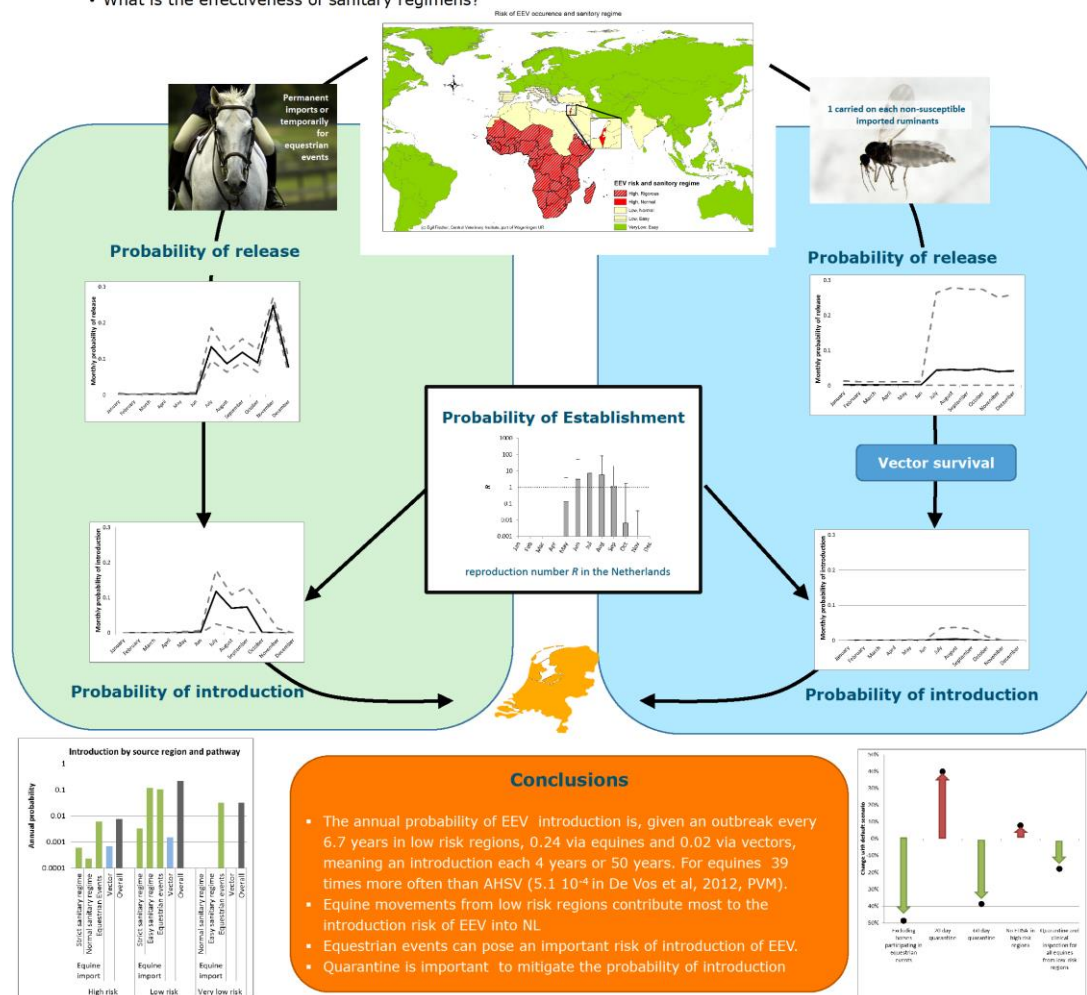
Evelyn Pamela Martínez López*, Céline Faverjon[§], Clazien J. de Vos*, Egil A.J. Fischer*[†]
^{*}Department of Epidemiology, Crisis organisation and Diagnostics, Central Veterinary Institute, part of Wageningen UR, The Netherlands [§] INRA UR346 Animal Epidemiology, VetagroSup, Marcy L'Etoile, France, ^p = presenting author



What is the probability of introduction of EEV in the Netherlands?

Equine encephalosis is a midge-borne disease of equines caused by EEV (Orbivirus, Reoviridae) exotic to the Netherlands. EEV is related to African horse sickness virus.

- What is the probability of EEV introduction into the Netherlands by either vectors or equines?
- What is the contribution of different source areas?
- What is the effectiveness of sanitary regimens?



Central Veterinary Institute, part of Wageningen UR
P.O. Box 65, 8200RA Lelystad, The Netherlands
Contact: e-mail: egil@egilfischer.nl
T + 31 (0)320 237380
www.wageningenUR.nl

Acknowledgements: We would like to thank Dr Klaas Frankena for his supervision of EPML at the QVE group of the Wageningen UR. This study was financially support through the EMIDA-ERANET project VICE by a visit of CF to the Central Veterinary Institute, part of Wageningen UR and the Ministry of Economic Affairs of the Netherlands (BO 20-009-009). EPML was supported by a scholarship of the Ecuadorian Secretaría Nacional de Educación Superior, Ciencia, Tecnología e Innovación.

Epidémiologie animale UR 346
INRA / VetAgro Sup
France
www.vetagro-sup.fr

Appendix 8: Poster presentation at the Journées de l'Ecole doctorale SVSAE in Clermont-Ferrand, 2015.

Quantitative Risk Analysis of Equine encephalosis virus (EEV) entry into France

Céline Favrejon¹, Pamela Martinez-Lopez², Agnes Leblond¹, Aline de Koeijer², Egil Fischer²

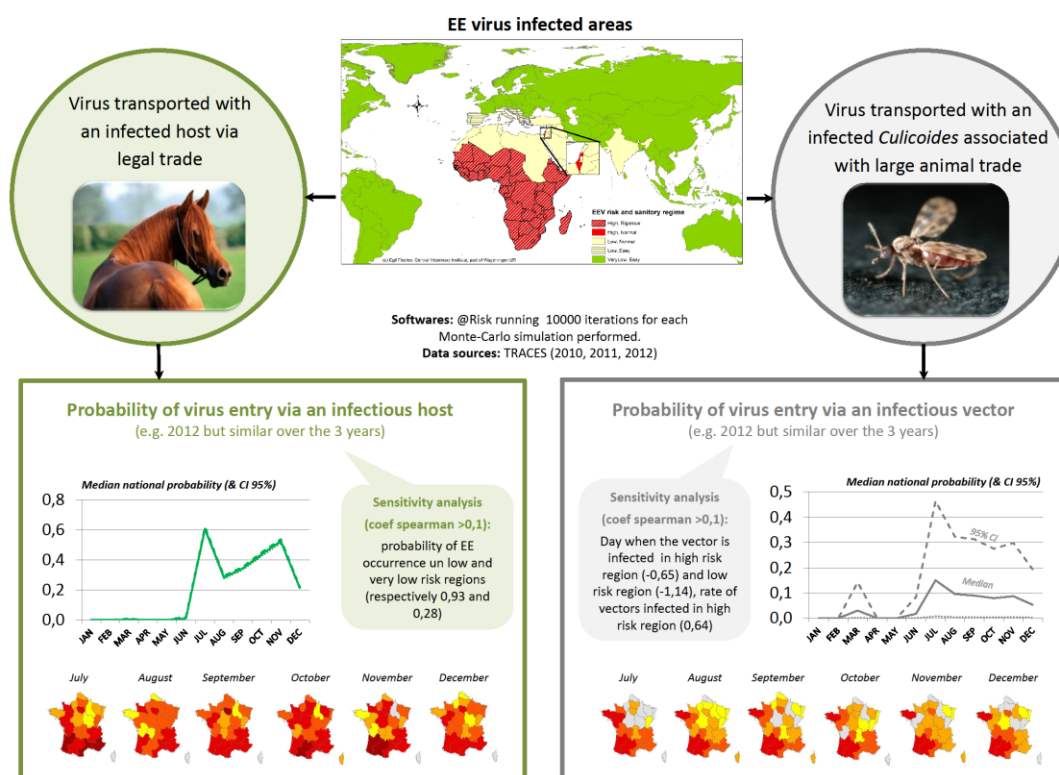
¹ INRA UR346 Animal Epidemiology, VetAgroSup, F-63280 Marcy l'Etoile, France : celine.favrejon@vetagro-sup.fr

² Department of Epidemiology, Crisis organisation and Diagnostics, Central Veterinary Institute, part of Wageningen UR, The Netherlands

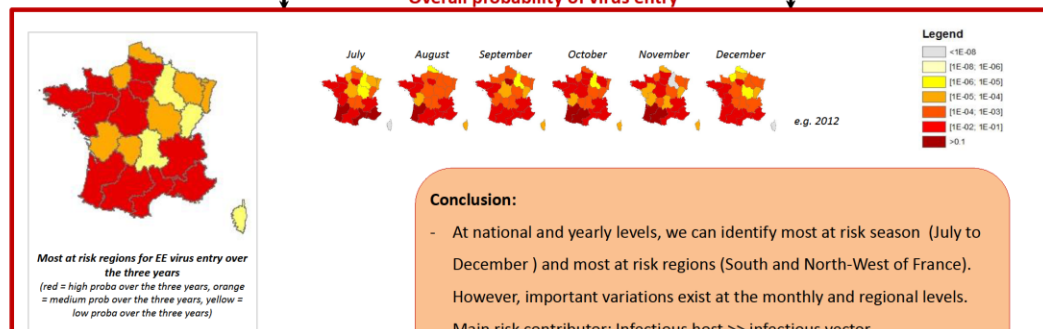


Equine encephalosis (EE) is a *Culicoides*-borne disease caused by an *Orbivirus* (*Reoviridae*) exotic to France and affecting all equidae. The disease is mainly asymptomatic but EE virus is related to African horse sickness virus, one of the most devastating disease for equine (mortality >90%), and to Bluetongue virus which was recently introduced and largely spread into Europe.

Risk of EE Virus introduction to France?



Overall probability of virus entry



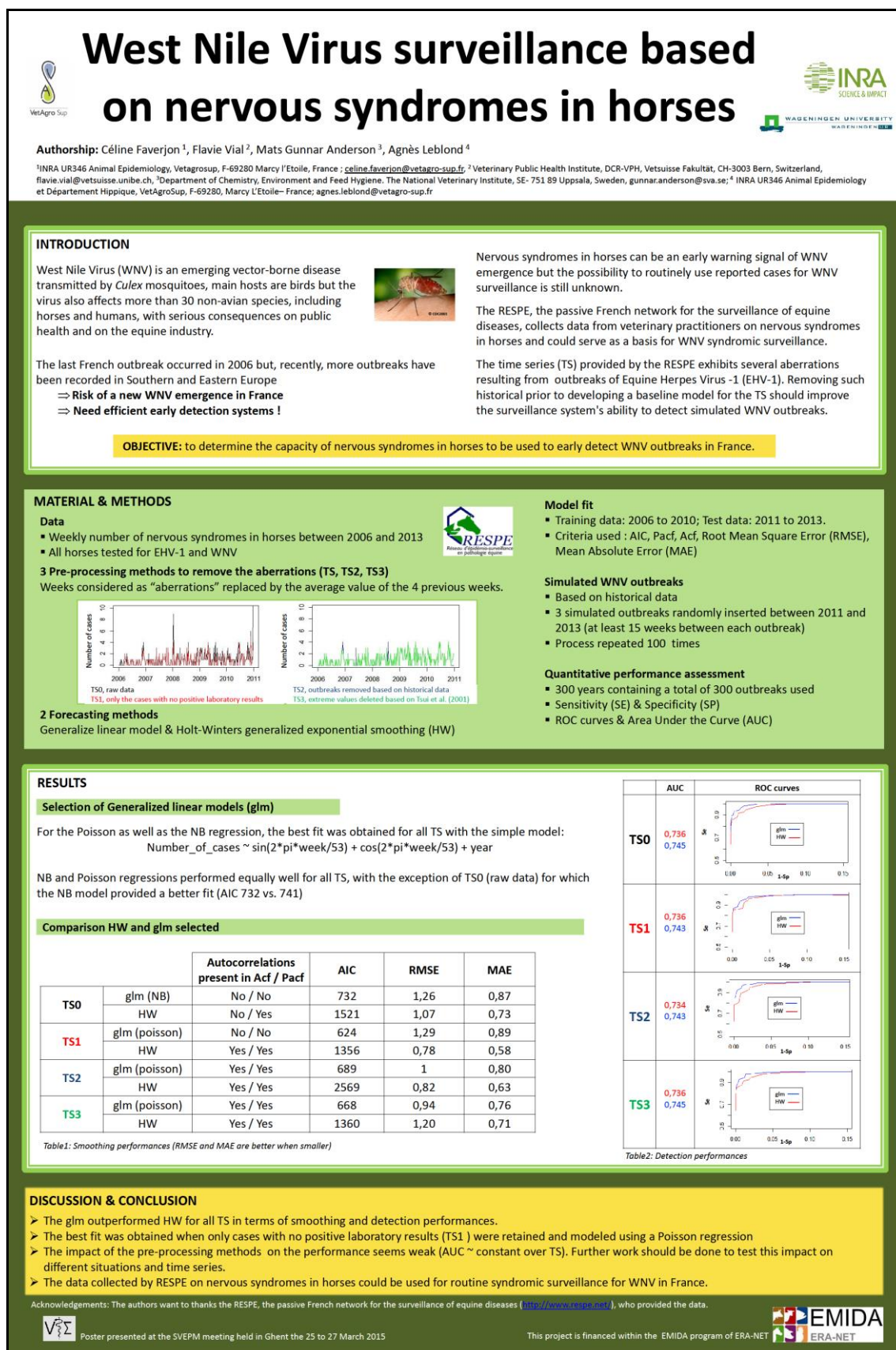
Acknowledgments: French Ministry of Agriculture, especially SIVEP and BICMA, who provided the data



This project is financed within the EMIDA program of ERA-NET

Poster presented at 'Journée de l'Ecole doctorale' SVSAE - May, 21st and 22nd, 2015

Appendix 9: Poster presentation at the Conference of the Society for Veterinary Epidemiology and Preventive Medicine (SVEPM) in Ghent, 2015.



Appendix 10: Article in a Professional journal published in 'Bulletin Epidémiologie et Santé Animale' by AEEMA. 2015. *Épidémiol. et santé anim.*, 2015, 67, 87-95

AIDE A LA DECISION EN SURVEILLANCE SYNDROMIQUE PAR LE CALCUL DE LA PROBABILITE DE CIRCULATION D'UN AGENT PATHOGENE *

Céline Faverjon¹, Mats Gunnar Andersson², Flavie Vial³,
Loïc Legrand^{4,5} et Agnès Leblond^{5,6}

RESUME

Dans ce travail nous proposons d'appliquer le cadre statistique utilisé lors de l'évaluation des pièces à conviction au cas de la surveillance syndromique. Nous avons utilisé cet outil pour évaluer les résultats de la surveillance syndromique en tant que « preuves » circonstanciées d'une épizootie. L'idée de base est d'exploiter les distributions de cas attendues afin de calculer le ratio entre la probabilité d'observer n cas lorsqu'une épizootie est en cours et la probabilité d'observer ce même nombre de cas lorsqu'il n'y a pas d'épizootie. Le rapport de vraisemblance ainsi obtenu correspond à la valeur de la preuve. D'après la règle de Bayes, en multipliant ce rapport avec les probabilités *a priori* que la maladie circule, on obtient une connaissance *a posteriori* sur le statut de la maladie. Cette approche a été appliquée à des séries temporelles représentant le nombre de chevaux présentant des symptômes nerveux. La séparation claire entre connaissance *a priori* et évaluation de la valeur de la preuve permet un raisonnement transparent apte à s'intégrer dans un processus de décision. Par ailleurs, l'approche bayésienne permet d'intégrer facilement les données de surveillance syndromique avec d'autres sources d'information telles que des évaluations de risque d'introduction et/ou de transmission d'agents pathogènes.

Mots-clés : Bayes, surveillance syndromique, West Nile

ABSTRACT

In this work we propose the adoption of a statistical framework to be used in the evaluation of forensic evidence as a tool for evaluating and presenting circumstantial "evidence" of a disease outbreak from syndromic surveillance. The basic idea is to exploit the predicted distributions of reported cases to calculate the ratio of the likelihood of observing n cases given an ongoing outbreak over the likelihood of observing n cases given no outbreak. The likelihood ratio defines the Value of Evidence. Using Bayes' rule, the prior odds for an ongoing outbreak are multiplied with V to obtain the posterior odds.

.../..

* Texte de la communication orale présentée au cours de la Journée scientifique AEEMA, 20 mars 2015

¹ INRA UR346 Épidémiologie animale, VetagroSup, F-69280 Marcy l'Etoile, France ; celine.faverjon@vetagro-sup.fr

² Department of Chemistry, Environment and Feed Hygiene. The National Veterinary Institute, SE- 751 89 Uppsala, Sweden

³ Veterinary Public Health Institute, DCR-VPH, Vetsuisse Fakultät, CH-3003 Bern, Switzerland

⁴ LABÉO - Frank Duncombe, Unité Risques Microbiens (U2RM), EA 4655, Normandie Université, Caen, France

⁵ Réseau d'EpidémioSurveillance en Pathologie Equine (RESPE), Caen, France.

⁶ INRA UR346 Épidémiologie animale et Département hippique, VetAgroSup, F-69280, Marcy L'Etoile, France

.../..

This approach was applied to time series on the number of horses showing nervous symptoms. The separation between prior beliefs about the probability of an outbreak and the strength of evidence from syndromic surveillance offers a transparent rational process suitable for supporting decision making. Furthermore, a Bayesian approach makes it possible to combine data from syndromic surveillance with results from predictive modeling and with information from other sources such as assessments of risks of disease introduction.

Keywords: Bayes, syndromic surveillance, West Nile



I - INTRODUCTION

1. RAPPELS SUR LA SURVEILLANCE SYNDROMIQUE

La surveillance syndromique est un concept apparu à la fin des années 90 et est devenue de plus en plus populaire en santé humaine mais également ces dernières années en santé animale [Dórea *et al.*, 2011]. Il n'existe à ce jour pas de définition unique et communément acceptée de la surveillance syndromique. Cependant, en général, la surveillance syndromique utilise des données pré-diagnostiques, souvent peu spécifiques, collectées en routine et analysées en temps réel [Katz *et al.*, 2011]. La surveillance syndromique a ainsi pour vocation de détecter précocement des maladies connues, comme la grippe saisonnière humaine (Hiller *et al.* 2013) ; (Ginsberg *et al.* 2009) ou inconnues, comme les attaques bioterroristes (Buehler *et al.* 2003). La surveillance syndromique ne remplace pas les approches traditionnelles de surveillance des maladies mais elle est un outil complémentaire intéressant de par sa rapidité, sa flexibilité et son bon rapport coût-bénéfice.

Les approches actuellement utilisées en surveillance syndromique cherchent d'abord à définir les propriétés normales de la série de données considérée lorsqu'aucun foyer de maladie n'est enregistré. L'objectif est de pouvoir ensuite détecter des événements anormaux tels que des épidémies ou des épizooties. Les méthodes de détection traditionnelles produisent une alarme lorsque les données observées dépassent les valeurs attendues en l'absence d'épidémie. Les algorithmes utilisés définissent ainsi un seuil

épidémique et fournissent une réponse finale de type oui/non : « non, aucune épidémie en cours », « oui, un événement inhabituel est en cours ».

Cette vision binaire d'une situation épidémiologique est simple mais elle n'est pas toujours suffisante et peut être compliquée à interpréter notamment lorsque les résultats de l'analyse sont dans une zone « grise » proche du seuil épidémique. De plus, ces résultats qualitatifs binaires sont également difficiles à combiner avec d'autres connaissances épidémiologiques, tels que le risque d'introduction ou la saisonnalité d'une maladie, qui entrent pourtant en compte lorsqu'une décision d'intervenir (ou non) doit être prise suite à la production d'une alarme. Développer des méthodes quantitatives transparentes, plus spécifiques, et facilement utilisables dans un processus de décision s'avère ainsi être un domaine de recherche prometteur.

2. OBJECTIFS DE L'ÉTUDE

L'objectif de ce travail est de tester l'applicabilité du cadre statistique bayésien pour la détection précoce d'épidémies en surveillance syndromique. Ces approches sont déjà utilisées notamment pour l'évaluation des pièces à convictions dans un cadre juridique (Foreman *et al.* 2003)(Drygajlo *et al.* 2003) ; (Morrison 2012) ; (Taroni *et al.* 2006)].

Nous détaillerons dans un premier temps le cadre théorique de la méthode développée, puis nous présenterons une application concrète de la méthode au cas de la surveillance de la fièvre de West Nile (WN). Le virus West Nile est un arbovirus

du genre *Flavivirus* (famille *Flaviviridae*). Il est principalement transmis par des moustiques du genre *Culex* (famille *Culicidae*). Les hôtes principaux sont les oiseaux mais le virus peut également affecter l'homme et plus de 30 espèces animales parmi lesquelles les équidés. Chez les personnes comme chez les chevaux, le virus peut provoquer des encéphalites mortelles et la présence de la maladie dans un territoire a ainsi des conséquences importantes en termes de santé publique et sur la filière équine.

Le virus a été identifié en Europe pour la première fois dans les années 60 et depuis, de nombreuses émergences ont été rapportées un peu partout sur le continent (Calistri et al. 2010). Même si le virus

est actuellement considéré comme endémique dans une grande partie de l'Europe (Ozdenerol et al. 2013), de plus en plus de cas sont rapportés ces dernières années dans le sud-est de l'Europe (exemple : Italie, Grèce, Bulgarie, Croatie, Serbie, Albanie) (Di Sabatino et al. 2014). L'impact de la maladie est plus limité en Europe que ce qui est observé en Amérique du Nord. Cependant, le nombre croissant d'émergences associé à la récente introduction de la lignée 2 en Europe font du virus WN une menace d'intérêt constant pour les pays européens (Hernández-Triana et al. 2014) ; (Bakonyi et al. 2006)(Calzolari et al. 2013). Développer des méthodes innovantes pour améliorer la détection d'émergences de WN s'avère ainsi particulièrement intéressant.

II - MÉTHODE : CADRE THÉORIQUE

1. CADRE STATISTIQUE BAYÉSIEN

La formule de Bayes se définit initialement comme suit :

$$\frac{P(M+)}{P(M-)} \times \frac{P(n|M+)}{P(n|M-)} = \frac{P(M+|n)}{P(M-|n)}$$

Appliquée au cas de la surveillance syndromique, on peut définir $M+$ comme le fait que la maladie soit présente, $M-$ comme le fait que la maladie soit absente et n comme le nombre de syndromes observés à un moment t .

Le premier terme $\frac{P(M+)}{P(M-)}$ est un odds qui correspond aux connaissances *a priori* qu'on a du statut présence vs absence de la maladie dans un territoire à un moment t . Il tient notamment compte des facteurs de risque connus tels que la saison ou la circulation du pathogène dans une région voisine. Il sera noté O_{pri} dans les paragraphes suivants.

Le second terme $\frac{P(n|M+)}{P(n|M-)}$ représente le rapport de deux probabilités. $p(n|M+)$ est la probabilité d'observer n cas de syndromes lorsque la maladie est présente et $p(n|M-)$ correspond à la probabilité d'observer le même nombre de syndromes lorsque la maladie est absente. Ce second terme

correspond au rapport de vraisemblance noté V dans les paragraphes suivants.

La combinaison des connaissances *a priori* sur la maladie avec les informations obtenues à partir des observations de terrain permet d'obtenir le troisième terme $\frac{P(M+|n)}{P(M-|n)}$. Cet odds correspond à la connaissance *a posteriori* sur le statut de la maladie dans le territoire. Il sera noté O_{post} dans les paragraphes suivants.

2. INTÉGRATION DANS UN PROCESSUS DE DÉCISION

Le rapport de vraisemblance O_{post} construit à partir du cadre statistique bayésien donne une appréciation quantitative de la probabilité de circulation de la maladie. Cependant, il est important de pouvoir déterminer à partir de quel seuil de probabilité il est utile (ou non) de déclencher une alarme et de mettre en place des mesures de contrôles et/ou d'investigation de la maladie.

La théorie de la décision évalue comment s'effectue une prise de décision rationnelle en présence d'incertitude (Gittelsohn 2013). La difficulté ici vient du fait que le résultat de l'option choisie est déterminé par l'élément incertain qui

ne peut être connu (et ce parfois même après que la décision ait été prise). Pour prendre une décision, il est donc nécessaire d'évaluer les conséquences de chaque action (en termes d'utilité ou de pertes) et leur probabilité de se produire. Le choix rationnel est celui pour lequel l'utilité est maximale c'est-à-dire, celui comportant les pertes attendues les plus faibles.

Ainsi, l'utilité d'une action A (noté $U(A)$) se définit comme la somme des utilités de l'action A dans la situation i (noté chacune $U(A_i)$), multipliée par la probabilité estimée d'être effectivement dans la situation i (noté $p(A_i)$) :

$$U(A) = \sum_{i=0}^n U(A_i) \times p(A_i)$$

Nous avons appliqué ce raisonnement au cas de la prise de décision en surveillance syndromique. Ici seuls deux types d'action sont possibles : $A1$, mise en œuvre de mesures de contrôle de la maladie et $A0$, pas d'action de contrôle. Leurs utilités respectives sont évaluées en fonction du contexte épidémiologique : soit $M1$ la maladie est présente, soit $M0$ la maladie est absente. Les pertes relatives

à chaque couple $U(A_i, M_j)$ sont fondées sur des évaluations socio-économiques. Les connaissances *a priori* sur le contexte sont fournies par O_{post} .

Ainsi, l'utilité de mettre en place des actions de contrôle ($U(A1)$) et l'utilité de ne rien faire ($U(A0)$) peuvent être définies comme suit :

$$U(A1) = (U(A1, M1) + U(A1, M0)) \times \frac{O_{post}}{1 + O_{post}}$$

$$U(A0) = (U(A0, M1) + U(A0, M0)) \times \frac{O_{post}}{1 + O_{post}}$$

Les valeurs de O_{post} pour lesquelles il sera utile d'agir plutôt que de ne rien faire, sont les valeurs situées au-dessus du seuil O_{post}^* défini pour $U(A1) = U(A0)$.

III - CAS PRATIQUE : SURVEILLANCE DE LA FIÈVRE DE WEST NILE

Ce cas propose d'utiliser la méthode développée afin de détecter les émergences de virus WN en France. Nous utiliserons pour cet exemple les données de syndromes nerveux chez les chevaux qui sont des indicateurs intéressants d'émergence de la maladie (Leblond et al. 2007). Les données de syndromes nerveux sont collectées en routine depuis 2006 par le réseau d'épidémiosurveillance en pathologie équine (RESPE).

1. CONNAISSANCES A PRIORI

La fièvre de WN est une maladie à transmission vectorielle. La probabilité de circulation du virus à des niveaux épizootiques est ainsi fortement liée à la dynamique de population de son vecteur principal, les moustiques du genre *Culex*. Ainsi les épizooties de WN présentent une saisonnalité marquée avec de nombreuses occurrences rapportées en été et à l'automne, des cas moins

fréquents au printemps et seulement quelques cas sporadiques en hiver.

En utilisant les données historiques relatives aux épizooties de WN rapportées en Europe ces dernières années, nous avons ainsi établi des O_{pri} de la maladie, différents selon les saisons. Les probabilités de circulation relatives au printemps, automne/été et hiver sont ainsi respectivement de 1:5:0,04.

2. RÉSULTATS DE LA SURVEILLANCE SYNDROMIQUE

2.1. SITUATION NON- EPIZOOTIQUE

Pour déterminer la probabilité d'observer un nombre de cas n lorsque la maladie est absente, $p(n|M-)$, nous avons utilisé les données collectées par le RESPE depuis 2006. Des traces d'épizooties

sont cependant présentes dans ces données et sont principalement dues aux émergences d'herpesvirus de type 1. Afin de modéliser au mieux la série de données en situation non-épidémiologique, nous les avons supprimées en utilisant uniquement les données provenant de chevaux qui n'ont pas obtenu de résultat de laboratoire positif. Les données de 2006 à 2010 ont été utilisées pour calibrer le modèle et les données de 2011 à 2012 pour le valider.

Plusieurs modèles ont été testés mais celui qui a obtenu les meilleures performances est celui utilisant une loi de poisson (AIC = 637,8 et GOF (déviante ajustée) = 1,156):

Nombre de cas $\sim \sin(2\pi \cdot t) + \cos(2\pi \cdot t) + \log(\text{histmean})$

où 'histmean' représente la moyenne des 53 semaines précédant t sans tenir compte des 10 semaines juste avant t afin de ne pas incorporer des données d'une épidémie éventuellement en cours.

2.2. SITUATION ÉPIDÉMIOLOGIQUE

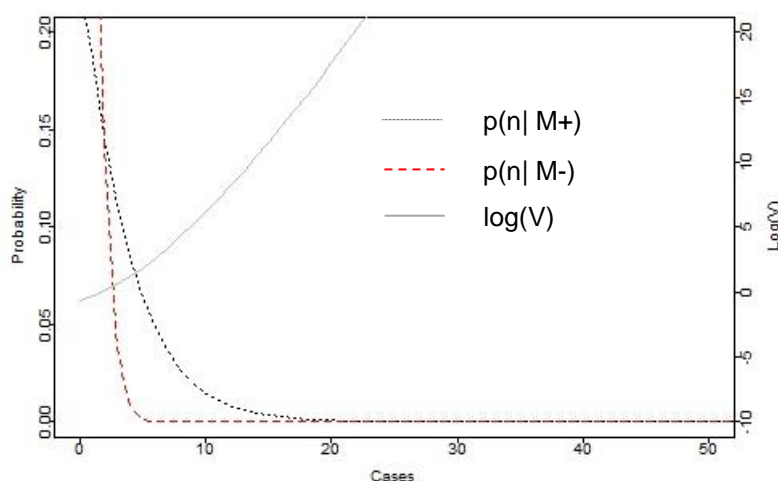
Afin de déterminer la probabilité d'observer un nombre de cas n lorsque la maladie est présente, $p(n|M+)$, nous avons utilisé les données trouvées dans la littérature sur des épidémies de WN ayant entraîné des symptômes nerveux chez les chevaux. Nous avons ainsi utilisé les nombres de symptômes rapportés chez les chevaux lors des émergences françaises de 2000 (Murgue et al. 2001) et 2004 (Leblond et al. 2007) et lors de l'épisode de WN en Italie en 1998 (Autorino et al. 2002).

Le nombre de cas observés en situation épidémiologique a été modélisé avec une loi négative binomiale de paramètre μ égal à 4,45 et θ égal à 0,94. Le nombre médian de cas attendus par semaine en cas d'épidémie a ainsi été estimé à 3 (IC 95 %: 0-18).

2.3. RAPPORT DE VRAISEMBLANCE

Le rapport de vraisemblance noté V correspond au rapport entre $p(n|M+)$ et $p(n|M-)$ et peut se représenter graphiquement comme présenté en figure 1.

Figure 1
Rapport de vraisemblance (V) et probabilités d'observer n cas de syndromes nerveux chez les chevaux lorsque la maladie est présente ($p(n|M+)$) et lorsque la maladie est absente ($p(n|M-)$)



3. SCÉNARIO DE DÉCISION

Afin de déterminer un seuil d'action, nous avons estimé des coûts fictifs mais proportionnels à leurs contributions relatives attendues (tableau 1). L'utilité de ne rien faire lorsque la maladie est

absente ($U(A0,M0)$) entraîne, quelle que soit la saison, des pertes nulles. L'utilité de ne rien faire lorsque la maladie est présente ($U(A0,M1)$) entraîne des coûts plus importants lors des périodes à risque à cause de l'impact de la maladie sur le tourisme et des coûts liés aux animaux

malades. L'utilité de mettre en place des mesures de contrôle alors que la maladie est absente ($U(A1, M0)$) entraîne quant à elle des pertes identiques quelle que soit la saison et liées uniquement aux actions mises en place pour contrôler la maladie (ex : démoustication, vaccination). L'utilité de mettre en place des

mesures de contrôle lorsque la maladie est effectivement présente ($U(A1, M1)$), entraîne, elle, des pertes liées aux actions de contrôle mises en place ainsi que des pertes indirectes et directes liées à la diffusion de la maladie qui sont estimées moins importantes que lorsqu'aucune mesure de contrôle n'est mise en places.

Tableau 1

Estimation du seuil d'action O_{post}^* à partir des couts estimés d'une émergence de WN en fonction des saisons.

A0 représente le fait de ne rien faire et A1 représente le fait de mettre en place des mesures de contrôle.

M0 correspond au fait que la maladie soit absente et M1 que la maladie soit présente.

Utilités	Automne/Été	Printemps	Hiver
$U(A0, M0)$	0	0	0
$U(A1, M0)$	-0.5 M€	-0.5 M€	-0.5 M€
$U(A0, M1)$	-5.1 M€	-10,1 M€	-5,3 M€
$U(A1, M1)$	-3.9 M€	-6,3 M€	-4,1 M€
$\text{Log}_{10}(O_{post}^*)$	-0,38	-0,88	-0,38

4. ÉVALUATION ET MISE EN ŒUVRE DU SYSTÈME

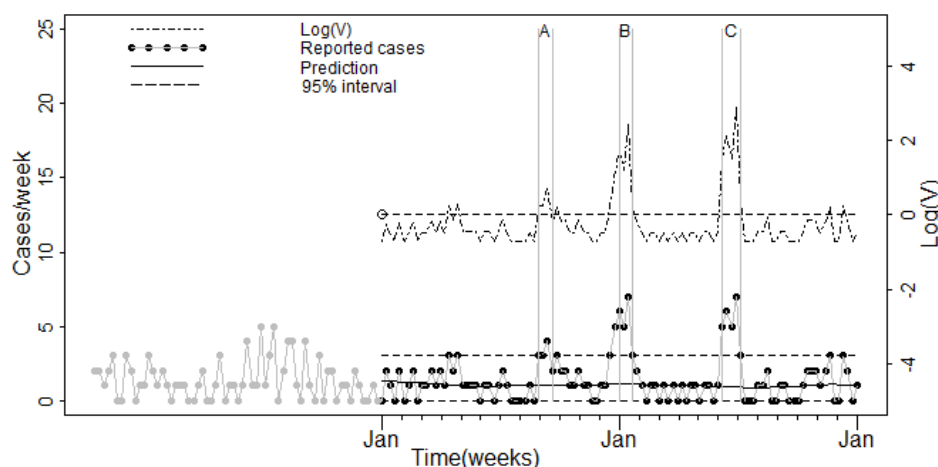
Les données de 2011 à 2012 du RESPE ont été utilisées pour détecter les épizooties. Afin d'illustrer notre méthode, trois épizooties ont été arbitrairement insérées dans les données : A.

émergence en automne d'un petit nombre de cas, B. émergence de taille « moyenne » en hiver et C. émergence « importante » au printemps. Pour chaque semaine de 2011 à 2012 ; le rapport de vraisemblance V a été calculé en fonction des données observées et attendues (cf. figure 2).

Figure 2

Données de syndromes nerveux des chevaux et calcul du rapport de vraisemblance pour chaque semaine.

Trois épizooties fictives insérées : A. insertion entre les semaines 36 et 39 ; B. insertion entre les semaines 1 et 4 ; C. insertion entre les semaines 24 et 28.



Un détail des cas observés par épizootie et des rapports de vraisemblance associés est présenté dans le tableau 2. Combiné avec les connaissances *a priori* sur la saisonnalité de la maladie (O_{pri}) défini précédemment, nos connaissances *a posteriori* (O_{post}) sur la maladie sont déduites. Combinée avec le seuil d'action défini à partir des couts estimés, la

valeur seuil du nombre de cas à partir de laquelle il est utile d'intervenir est déduite. Ainsi, en automne/été, quatre cas de syndromes nerveux suffisent à déclencher une action alors qu'en hiver, sept cas de syndromes nerveux observés ne déclenchent toujours pas d'action vis-à-vis d'une alerte WN.

Tableau 2
Détails des épizooties fictives insérées et valeurs seuils pour le déclenchement d'une action en fonction du nombre de syndromes observés et de la saison

	A. Automne/Été		B. Printemps		B. Hiver	
$\text{Log}_{10}(O_{pri})$	-0,99		-1,78		-3,03	
Nombre de cas observés	3	4	5	7	5	7
$\text{Log}_{10}(V)$	0,26	0,74	1,54	3,11	1,24	2,6
$\text{Log}_{10}(O_{post})$	-0,73	-0,25	-0,24	1,33	-1,77	-0,43
Seuil d'Action $\text{Log}_{10}(O_{post}^*)$	-0,38		-0,88		-0,38	
Alerte? $O_{post} > O_{post}^*$	Non	Oui	Oui	Oui	Non	Non

IV - DISCUSSION

Dans ce travail, nous avons montré comment le cadre statistique bayésien peut être incorporé dans un système d'aide à la décision en surveillance syndromique et comment il peut être utilisé pour l'évaluation des risques et la prise de décision éclairée. L'approche proposée est en accord avec le guide d'analyse de risque (OIE 2010) puisqu'elle permet une séparation explicite des hypothèses (O_{pri}), des preuves scientifiques (V) et des critères pour la prise de décision ($U(A0)$ et $U(A1)$). Par ailleurs, la manière dont les preuves scientifiques sont évaluées est également transparente et quantitative, ce qui limite les interprétations subjectives. Les critères utilisés pour la prise de décision sont également en accord avec ce qui est préconisé par l'OIE. En effet, le code sanitaire pour les animaux terrestres recommande de fonder la décision de mise en place de mesures de contrôle sur l'équilibre entre le coût des activités de lutte et les conséquences économiques et sociales d'une réponse retardée.

L'approche utilisée permet de combiner facilement les données de surveillance syndromique avec d'autres sources d'information épidémiologiques. Dans l'exemple développé ici, la prise en compte

des données de saisonnalité dans l'interprétation des résultats de la surveillance syndromique permet ainsi de déclencher des alertes pour la fièvre de WN préférentiellement dans les périodes à haut risque. La détection des émergences est ainsi plus sensible en période à haut risque et plus spécifique en période à faible risque. Par exemple, en automne/été, quatre cas de syndromes nerveux suffisent à déclencher une alerte pour WN alors qu'en hiver sept cas de syndromes nerveux observés ne déclenchent pas d'alerte pour WN. Notre approche permet donc de rendre l'interprétation des résultats de la surveillance syndromique plus spécifique sans pour autant diminuer la sensibilité. Ceci est particulièrement intéressant lorsqu'il faut détecter des maladies connues.

Dans l'exemple développé ici, seules les connaissances sur la saisonnalité de la circulation du virus WN ont été intégrées. Cependant, des données quantitatives beaucoup plus complexes peuvent être utilisées comme par exemple les résultats de modèles de risque d'introduction ou de risque de diffusion. De même, nous avons travaillé uniquement à une échelle temporelle

mais le modèle pourrait être adapté pour détecter des agrégats spatio-temporels de cas. Cette approche offre ainsi des perspectives d'utilisation et de développement ultérieurs intéressants en permettant de passer d'informations épidémiologiques morcelées à une vision plus synthétique et intégrée.

Si la méthode offre des perspectives prometteuses, certaines limites sont cependant à considérer. Le cadre statistique bayésien a ici été appliqué en estimant les probabilités *a priori* de manière probabiliste à partir des données disponibles et non par le biais d'inférence bayésienne. L'approche probabiliste est relativement simple à mettre en place et à présenter. Elle a cependant le défaut de ne pas complètement rendre compte de l'incertitude des hypothèses de départ. Des approches utilisant les méthodes d'inférence ont été proposées par plusieurs auteurs pour détecter des agrégats spatio-temporels de cas (Banks et al. 2012); (Zou et al. 2010). Cependant, si ces

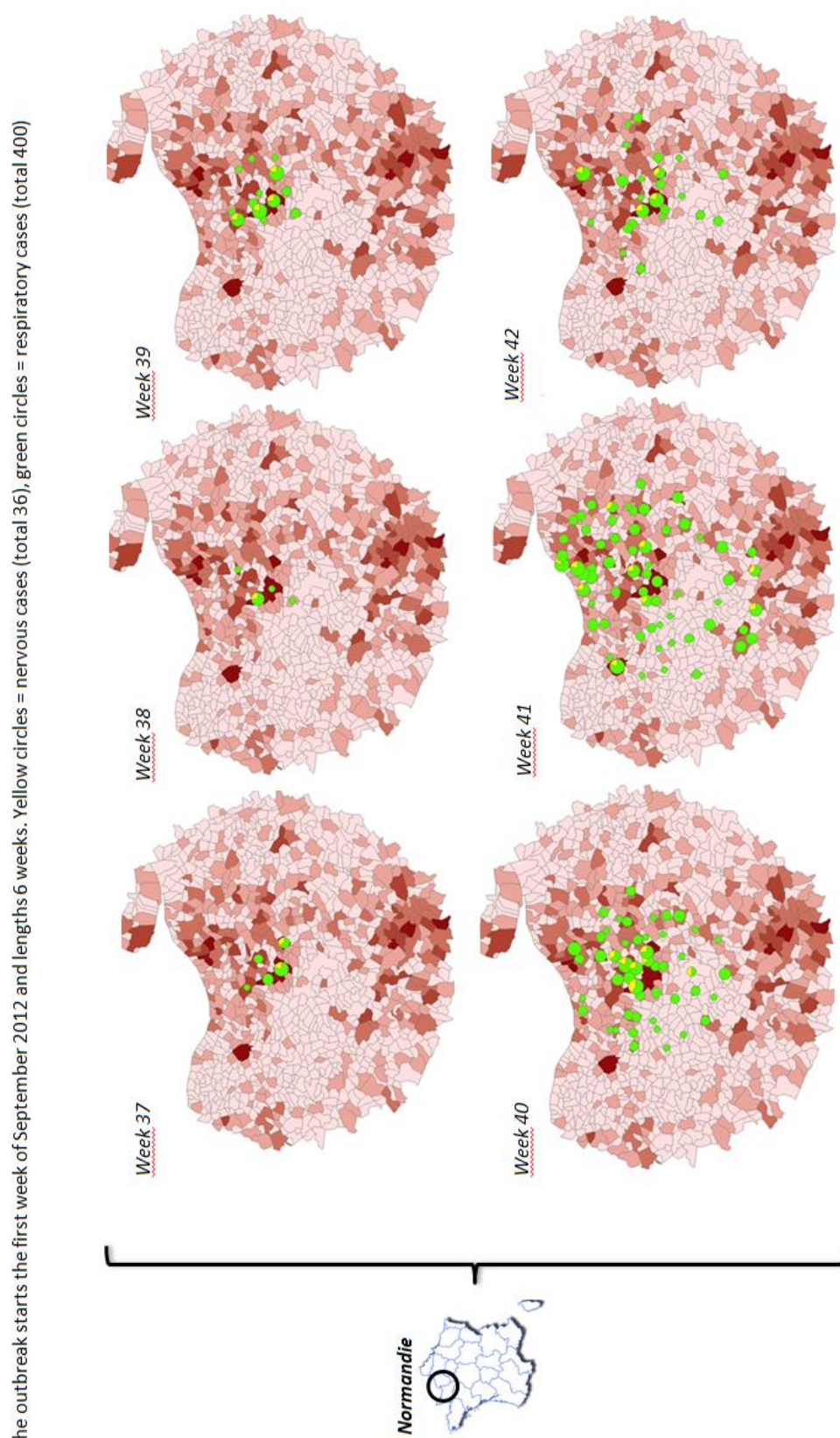
approches sont techniquement réalisables et plus robustes d'un point de vue mathématique, elles restent encore assez théoriques et compliquées à présenter à un public non familier des méthodes bayésiennes tel que les décideurs politiques (Banks et al. 2012).

Concernant l'exemple développé, il est important de noter que nous avons ici considéré que les semaines étaient indépendantes les unes des autres. Cette hypothèse simple a été choisie pour présenter au mieux la méthode et ne constitue pas une limite à son application. En effet, il est tout à fait possible de construire un système plus complexe considérant non pas les semaines de manière séparées mais les preuves cumulées sur plusieurs semaines. Par ailleurs, si un système de surveillance de WN fondé sur cette approche devait effectivement être implémenté, il faudrait nécessairement mieux préciser les estimations de nombre de cas utilisées, la saisonnalité des données ainsi que les estimations coûts-bénéfice.

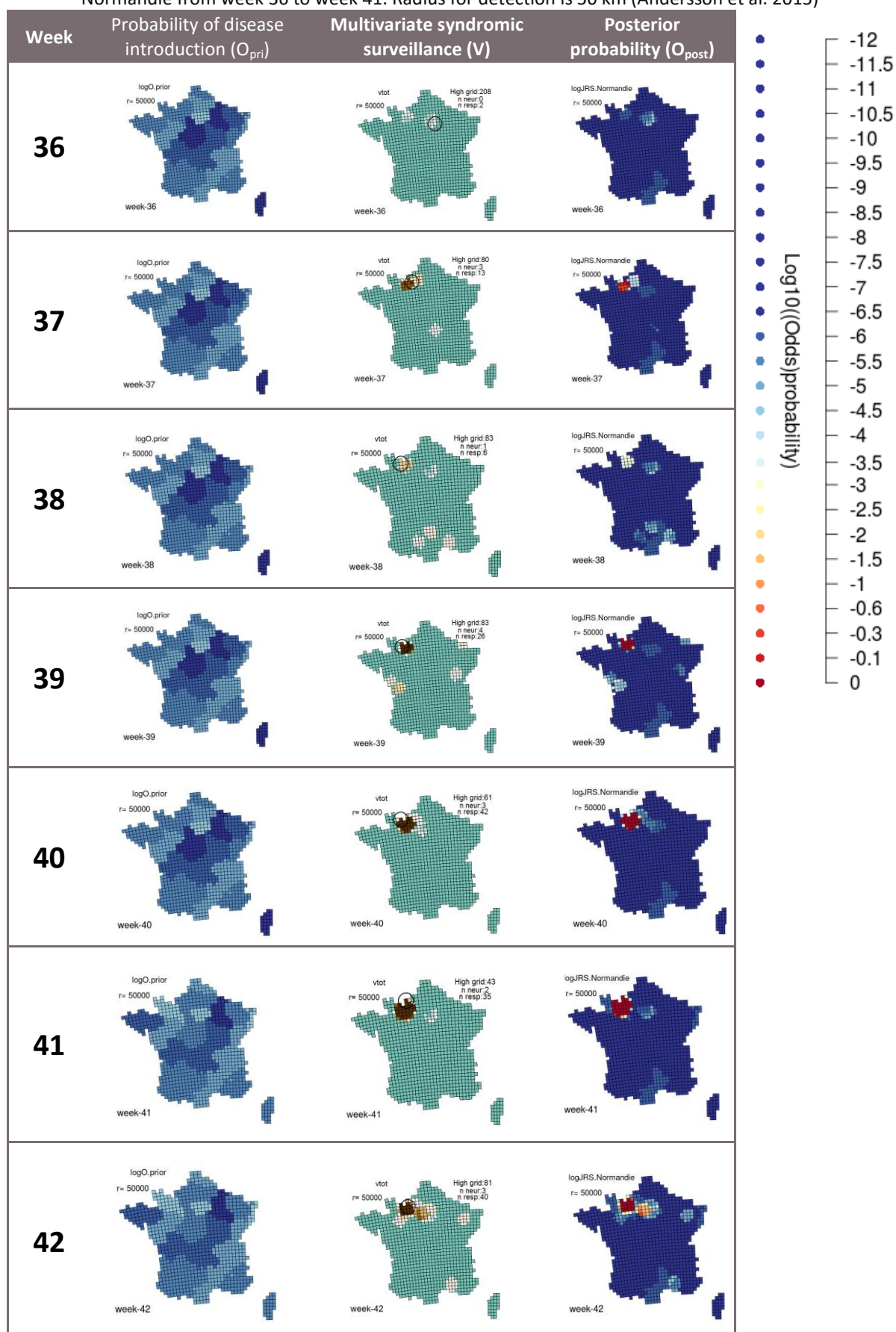
V - CONCLUSION

L'utilisation du cadre statistique bayésien en surveillance syndromique est donc une approche prometteuse pour l'amélioration de la surveillance des maladies connues. Elle permet de synthétiser de manière quantitative les diverses sources d'informations épidémiologiques disponibles et de les intégrer aisément dans un processus de décision rationnel prenant en compte des analyses coûts bénéfices. Des travaux ultérieurs pourront permettre de perfectionner la méthode et d'évaluer son intérêt dans des systèmes de surveillance complexes. L'ensemble des analyses détaillées est disponible dans l'article publié par Anderson, Faverjon *et al.* (Andersson et al. 2014).

Appendix 11: EEV simulated outbreak in Normandie.



Appendix 12: Posterior probability of EE outbreak in France in year 2012 during the simulated outbreak in Normandie from week 36 to week 41. Radius for detection is 50 km (Andersson et al. 2015)



REFERENCES

- Abdo-Salem, Shaif, Agnès Waret-Szkuta, François Roger, Marie-Marie Olive, Khalid Saeed, et Véronique Chevalier. 2011. « Risk Assessment of the Introduction of Rift Valley Fever from the Horn of Africa to Yemen via Legal Trade of Small Ruminants ». *Tropical Animal Health and Production* 43 (2): 471-80. doi:10.1007/s11250-010-9719-7.
- Abdullah, Abu S.M., Brian Tomlinson, Clive S. Cockram, et G. Neil Thomas. 2003. « Lessons from the Severe Acute Respiratory Syndrome Outbreak in Hong Kong ». *Emerging Infectious Diseases* 9 (9): 1042-45. doi:10.3201/eid0909.030366.
- Aharonson-Raz, Karin, Amir Steinman, Velizar Bumbarov, Sushila Maan, Narender Singh Maan, Kyriaki Nomikou, Carrie Batten, et al. 2011. « Isolation and Phylogenetic Grouping of Equine Encephalosis Virus in Israel ». *Emerging Infectious Diseases* 17 (10): 1883-86. doi:10.3201/eid1710.110350.
- Altizer, Sonia, Andrew Dobson, Parvize Hosseini, Peter Hudson, Mercedes Pascual, et Pejman Rohani. 2006. « Seasonality and the Dynamics of Infectious Diseases ». *Ecology Letters* 9 (4): 467-84. doi:10.1111/j.1461-0248.2005.00879.x.
- Andersson, Mats Gunnar, Céline Faverjon, Flavie Vial, Loïc Legrand, et Agnès Leblond. 2014. « Using Bayes' Rule to Define the Value of Evidence from Syndromic Surveillance ». *PloS One* 9 (11): e111335. doi:10.1371/journal.pone.0111335.
- Andersson, Mats Gunnar, Egil AJ Fischer, Céline Faverjon, Jörn Gerthmann, Maya Gussmann, Yves Van der Stede, P. Hopp, René Bødker, et Agnès Leblond. 2015. « A joint risk score method for risk-based surveillance of vector-borne animal diseases », Manuscript in preparation, .
- Anonymous. 2007. « Programme de surveillance vétérinaire de la fièvre West-Nile ». Note de service DGAL/SDSPA/N2007-8136. Direction générale de l'alimentation. <http://agriculture.gouv.fr/IMG/pdf/dgaln20078136z.pdf>.
- . 2011. « African Horse Sickness Outbreak Western Cape February - March 2011 ». 7. South Africa: Veterinary Services of the Western Cape Province. http://www.africanhorsesickness.co.za/Documents%5Cdoc_47.pdf.
- Asseged, B.D., T. Habtemariam, B. Tameru, et D. Nganwa. 2012. « The risk of introduction of equine infectious anemia virus into USA via cloned horse embryos imported from Canada ». *Theriogenology* 77 (2): 445-58. doi:10.1016/j.theriogenology.2011.08.019.
- Autorino, Gian Luca, Antonio Battisti, Vincent Deubel, Giancarlo Ferrari, Riccardo Forletta, Armando Giovannini, Rossella Lelli, Severine Murri, et Maria Teresa Scicluna. 2002. « West Nile virus Epidemic in Horses, Tuscany Region, Italy ». *Emerging Infectious Diseases* 8 (12): 1372-78. doi:10.3201/eid0812.020234.
- Babo Martins, S, et Jonathan Rushton. 2014. « Cost-effectiveness analysis: adding value to assessment of animal health, welfare and production ». *Scientific and Technical Review* 33 (3). http://www.oie.int/fileadmin/Home/eng/Publications_%26_Documentation/docs/pdf/revue_plurithematique/2014/22012014-00025-EN-BaboMartins.pdf.
- Backer, Jantien A., et Gonnie Nodelijk. 2011. « Transmission and Control of African Horse Sickness in The Netherlands: A Model Analysis ». *PLoS ONE* 6 (8). doi:10.1371/journal.pone.0023066.
- Bakonyi, Tamás, Emőke Ferenczi, Károly Erdélyi, Orsolya Kutasi, Tibor Csörgő, Bernhard Seidel, Herbert Weissenböck, Katharina Brugger, Enikő Bán, et Norbert Nowotny. 2013. « Explosive spread of a neuroinvasive lineage 2 West Nile virus in Central Europe, 2008/2009 ». *Veterinary Microbiology*, Special Issue: One World, One Health, One Virology, 165 (1–2): 61-70. doi:10.1016/j.vetmic.2013.03.005.

- Bakonyi, Tamás, Éva Ivanics, Károly Erdélyi, Krisztina Ursu, Emőke Ferenczi, Herbert Weissenböck, et Norbert Nowotny. 2006. « Lineage 1 and 2 Strains of Encephalitic West Nile Virus, Central Europe ». *Emerging Infectious Diseases* 12 (4): 618-23. doi:10.3201/eid1204.051379.
- Banks, David, Gauri Datta, Alan Karr, James Lynch, Jarad Niemi, et Francisco Vera. 2012. « Bayesian CAR models for syndromic surveillance on multiple data streams: Theory and practice ». *Information Fusion*, Special Issue on Information Fusion Applications to Human Health and Safety, 13 (2): 105-16. doi:10.1016/j.inffus.2009.10.005.
- Barnard, B J. 1993. « Circulation of African horsesickness virus in zebra (*Equus burchelli*) in the Kruger National Park, South Africa, as measured by the prevalence of type specific antibodies ». *The Onderstepoort journal of veterinary research* 60 (2): 111-17.
- Baylis, M, L O'Connell, et P S Mellor. 2008. « Rates of bluetongue virus transmission between *Culicoides sonorensis* and sheep ». *Medical and veterinary entomology* 22 (3): 228-37. doi:10.1111/j.1365-2915.2008.00732.x.
- Beck, Cécile, Miguel Jimenez-Clavero, Agnès Leblond, Benoît Durand, Norbert Nowotny, Isabelle Leparco-Goffart, Stéphan Zientara, Elsa Jourdain, et Sylvie Lecollinet. 2013. « Flaviviruses in Europe: Complex Circulation Patterns and Their Consequences for the Diagnosis and Control of West Nile Disease ». *International Journal of Environmental Research and Public Health* 10 (11): 6049-83. doi:10.3390/ijerph10116049.
- Bellini, R., M. Calzolari, A. Mattivi, M. Tamba, P. Angelini, P. Bonilauri, A. Albieri, et al. 2014. « The Experience of West Nile Virus Integrated Surveillance System in the Emilia-Romagna Region: Five Years of Implementation, Italy, 2009 to 2013 ». *Euro Surveillance* 19 (44).
- Bellini, Romeo, Herve Zeller, et Wim Van Bortel. 2014. « A Review of the Vector Management Methods to Prevent and Control Outbreaks of West Nile Virus Infection and the Challenge for Europe ». *Parasites & Vectors* 7 (1): 323. doi:10.1186/1756-3305-7-323.
- Bessell, P. R., R. A. Robinson, N. Golding, K. R. Searle, I. G. Handel, L. A. Boden, B. V. Purse, et B. M. de C. Bronsvoort. 2014. « Quantifying the Risk of Introduction of West Nile Virus into Great Britain by Migrating Passerine Birds ». *Transboundary and Emerging Diseases*. doi:10.1111/tbed.12310.
- Beugnet, Frédéric, Karine Chalvet-Monfray, et Harilaos Loukos. 2009. « FleaTickRisk: a meteorological model developed to monitor and predict the activity and density of three tick species and the cat flea in Europe ». *Geospatial health* 4 (1): 97-113.
- Box, George E. P., Gwilym M. Jenkins, et Gregory C. Reinsel. 2008. *Time Series Analysis: Forecasting & Control*. 4^e éd. Hoboken, N.J: Wiley.
- Bozdogan, Hamparsum. 1987. « Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions ». *Psychometrika* 52 (3): 345-70.
- Bronsvoort, B.M. de C., L. Alban, et M. Greiner. 2008. « Quantitative assessment of the likelihood of the introduction of classical swine fever virus into the Danish swine population ». *Preventive Veterinary Medicine* 85 (3-4): 226-40. doi:10.1016/j.prevetmed.2008.01.013.
- Brown, Eleanor B. E., Amie Adkin, Anthony R. Fooks, Ben Stephenson, Jolyon M. Medlock, et Emma L. Snary. 2012. « Assessing the Risks of West Nile Virus-Infected Mosquitoes from Transatlantic Aircraft: Implications for Disease Emergence in the United Kingdom ». *Vector-Borne and Zoonotic Diseases* 12 (4): 310-20. doi:10.1089/vbz.2010.0176.
- Brown, G. 2012. « California Mosquito-borne Virus Surveillance & Response Plan ». Davis: Mosquito & Vector control Association of California, University of California at Davis. <http://www.cdph.ca.gov/HealthInfo/discond/Documents/CAResponsePlanMay2012.pdf>.

- Brownson, Ross C., Elizabeth A. Baker, Terry L. Leet, Kathleen N. Gillespie, et William R. True. 2010. *Evidence-Based Public Health*. USA: Oxford University Press.
- Brownson, Ross C., et Ellen Jones. 2009. « Bridging the gap: Translating research into policy and practice ». *Preventive Medicine* 49 (4): 313-15. doi:10.1016/j.ypmed.2009.06.008.
- Brugger, Katharina, et Franz Rubel. 2013. « Bluetongue Disease Risk Assessment Based on Observed and Projected *Culicoides obsoletus* spp. Vector Densities ». Édité par Simon Gubbins. *PLoS ONE* 8 (4): e60330. doi:10.1371/journal.pone.0060330.
- Buckeridge, David L. 2007. « Outbreak Detection through Automated Surveillance: A Review of the Determinants of Detection ». *Journal of Biomedical Informatics* 40 (4): 370-79. doi:10.1016/j.jbi.2006.09.003.
- Buckeridge, David L., Howard Burkom, Murray Campbell, William R. Hogan, et Andrew W. Moore. 2005. « Algorithms for rapid outbreak detection: a research synthesis ». *Journal of Biomedical Informatics* 38 (2): 99-113. doi:10.1016/j.jbi.2004.11.007.
- Buckeridge, David L., Anna Okhmatovskaia, Samson Tu, Martin O'Connor, Csongor Nyulas, et Mark A. Musen. 2008. « Understanding Detection Performance in Public Health Surveillance: Modeling Aberrancy-detection Algorithms ». *Journal of the American Medical Informatics Association : JAMIA* 15 (6): 760-69. doi:10.1197/jamia.M2799.
- Buehler, James W., Ruth L. Berkelman, David M. Hartley, et Clarence J. Peters. 2003. « Syndromic Surveillance and Bioterrorism-related Epidemics ». *Emerging Infectious Diseases* 9 (10): 1197-1204. doi:10.3201/eid0910.030231.
- Bühler, Peter, Richard Hopkins, J. Marc Overhage, Daniel Sosin, et Van Tong. 2004. « Framework for Evaluating Public Health Surveillance Systems for Early Detection of Outbreaks ». 53. Morbidity and Mortality Weekly Report. CDC. http://link.springer.com/chapter/10.1007/978-3-642-18715-5_1.
- Bunning, M. L., R. A. Bowen, C. B. Cropp, K. G. Sullivan, B. S. Davis, N. Komar, M. S. Godsey, et al. 2002. « Experimental infection of horses with West Nile virus ». *Emerging infectious diseases* 8 (4): 380-86. doi:10.3201/eid0804.010239.
- Burkom, Howard S., Liane Ramac-Thomas, Steven Babin, Rekha Holtry, Zaruhi Mnatsakanyan, et Cynthia Yund. 2011. « An Integrated Approach for Fusion of Environmental and Human Health Data for Disease Surveillance ». *Statistics in Medicine* 30 (5): 470-79. doi:10.1002/sim.3976.
- Calistri, Paolo, Armando Giovannini, Zdenek Hubalek, Aurelia Ionescu, Federica Monaco, Giovanni Savini, et Rossella Lelli. 2010. « Epidemiology of West Nile in Europe and in the Mediterranean Basin ». *The Open Virology Journal* 4: 29-37. doi:10.2174/1874357901004020029.
- Calistri, P., L. Savini, L. Candeloro, D. Di Sabatino, F. Cito, R. Bruno, et M. L. Danzetta. 2014. « A Transitional Model for the Evaluation of West Nile Virus Transmission in Italy ». *Transboundary and Emerging Diseases*. doi:10.1111/tbed.12290.
- Calzolari, Mattia, Federica Monaco, Fabrizio Montarsi, Paolo Bonilauri, Silvia Ravagnan, Romeo Bellini, Giovanni Cattoli, et al. 2013. « New Incursions of West Nile Virus Lineage 2 in Italy in 2013: The Value of the Entomological Surveillance as Early Warning System ». *Veterinaria Italiana* 49 (3): 315-19. doi:10.12834/VetIt.1308.04.
- Campbell, Grant L., Anthony A. Marfin, Robert S. Lanciotti, et Duane J. Gubler. 2002. « West Nile Virus ». *The Lancet. Infectious Diseases* 2 (9): 519-29.
- Cantile, C., G. Di Guardo, C. Eleni, et M. Arispici. 2000. « Clinical and Neuropathological Features of West Nile Virus Equine Encephalomyelitis in Italy ». *Equine Veterinary Journal* 32 (1): 31-35. doi:10.2746/042516400777612080.

- Cardoen, Sabine, Xavier Van Huffel, Dirk Berkvens, Sophie Quoilin, Genevieve Ducoffre, Claude Saegerman, Niko Speybroeck, et al. 2009. « Evidence-based semiquantitative methodology for prioritization of foodborne zoonoses ». *Foodborne pathogens and disease* 6 (9): 1083-96. doi:10.1089/fpd.2009.0291.
- Carpenter, T. E. 2002. « Evaluation and Extension of the Cusum Technique with an Application to Salmonella Surveillance ». *Journal of Veterinary Diagnostic Investigation* 14 (3): 211-18. doi:10.1177/104063870201400304.
- Carpenter, Tim E. 2001. « Methods to investigate spatial and temporal clustering in veterinary epidemiology ». *Preventive Veterinary Medicine* 48 (4): 303-20. doi:10.1016/S0167-5877(00)00199-9.
- Castejón-Montijano, R., et L. Rodríguez-Fernández. 2011. « Economic Impact of the Horse Industry: A Special Reference to Spain ». *Journal of Agricultural Science & Technology* 1 (3A): 326.
- Castillo-Olivares, Javier, et James Wood. 2004. « West Nile virus infection of horses ». *Veterinary research* 35 (4): 467-83. doi:10.1051/vetres:2004022.
- Chaber, Anne-Lise, Sophie Allebone-Webb, Yves Lignereux, Andrew A. Cunningham, et J. Marcus Rowcliffe. 2010. « The Scale of Illegal Meat Importation from Africa to Europe via Paris ». *Conservation Letters* 3 (5): 317-21. doi:10.1111/j.1755-263X.2010.00121.x.
- Chaintoutis, Serafeim C., Chrysostomos I. Dovas, Maria Papanastassopoulou, Sandra Gewehr, Kostas Danis, Cécile Beck, Sylvie Lecollinet, et al. 2014. « Evaluation of a West Nile virus surveillance and early warning system in Greece, based on domestic pigeons ». *Comparative Immunology, Microbiology and Infectious Diseases* 37 (2): 131-41. doi:10.1016/j.cimid.2014.01.004.
- Chai, T., et R. R. Draxler. 2014. « Root Mean Square Error (RMSE) or Mean Absolute Error (MAE)? – Arguments against Avoiding RMSE in the Literature ». *Geoscientific Model Development* 7 (3): 1247-50. doi:10.5194/gmd-7-1247-2014.
- Chan, Ta-Chien, Chwan-Chuen King, Muh-Yong Yen, Po-Huang Chiang, Chao-Sheng Huang, et Chuhsing K. Hsiao. 2010. « Probabilistic Daily ILI Syndromic Surveillance with a Spatio-Temporal Bayesian Hierarchical Model ». *PLoS ONE* 5 (7): e11626. doi:10.1371/journal.pone.0011626.
- Chaskopoulou, Alexandra, Chrysostomos I. Dovas, Serafeim C. Chaintoutis, Javid Kashefi, Philip Koehler, et Maria Papanastassopoulou. 2013. « Detection and Early Warning of West Nile Virus Circulation in Central Macedonia, Greece, Using Sentinel Chickens and Mosquitoes ». *Vector-Borne and Zoonotic Diseases* 13 (10): 723-32. doi:10.1089/vbz.2012.1176.
- Chevalier, Veronique, Sylvie Lecollinet, et Benoit Durand. 2011. « West Nile Virus in Europe: A Comparison of Surveillance System Designs in a Changing Epidemiological Context ». *Vector Borne and Zoonotic Diseases* 11 (8): 1085-91. doi:10.1089/vbz.2010.0234.
- Choi, Byeong Yeob, Ho Kim, Un Yeong Go, Jong-Hyeon Jeong, et Jae Won Lee. 2010. « Comparison of Various Statistical Methods for Detecting Disease Outbreaks ». *Computational Statistics* 25 (4): 603-17. doi:10.1007/s00180-010-0191-7.
- Cochrane, A. L. 1999. *Effectiveness and efficiency: Random reflections on health services*. The Nuffield Provincial Hospitals Trust. Taylor & Francis.
http://www.nuffieldtrust.org.uk/sites/files/nuffield/publication/Effectiveness_and_Efficiency.pdf
- Codex Alimentarius. 1999. « Principles and guidelines for the conduct of microbiological risk management ». Guidelines CAC/GL-30. Codex Alimentarius.
http://www.codexalimentarius.org/input/download/standards/10741/CXG_063e.pdf.

- Colpitts, Tonya M., Michael J. Conway, Ruth R. Montgomery, et Erol Fikrig. 2012. « West Nile Virus: Biology, Transmission, and Human Infection ». *Clinical Microbiology Reviews* 25 (4): 635-48. doi:10.1128/CMR.00045-12.
- Commission Decision. 2003. *Commission Decision concerning the development of an integrated computerised veterinary system known as Traces*.
- Cox, Louis Anthony (Tony), Djangir Babayev, et William Huber. 2005. « Some Limitations of Qualitative Risk Rating Systems ». *Risk Analysis* 25 (3): 651-62. doi:10.1111/j.1539-6924.2005.00615.x.
- Crafford, J. E., A. J. Guthrie, M. Van Vuuren, P. P. C. Mertens, J. N. Burroughs, P. G. Howell, C. A. Batten, et C. Hamblin. 2011. « A competitive ELISA for the detection of group-specific antibody to equine encephalosis virus ». *Journal of Virological Methods* 174 (1-2): 60-64. doi:10.1016/j.jviromet.2011.03.024.
- Crafford, J. E., A. J. Guthrie, M. van Vuuren, P. P. C. Mertens, J. N. Burroughs, P. G. Howell, et C. Hamblin. 2003. « A group-specific, indirect sandwich ELISA for the detection of equine encephalosis virus antigen ». *Journal of Virological Methods* 112 (1-2): 129-35. doi:10.1016/S0166-0934(03)00202-7.
- Daix, Charlène. 2014. « Typologie des vétérinaires sentinelles ». *RESPE*. <http://www.respe.net/node/1925>.
- Damasio, Antonio R. 1995. *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Avon Books.
- Decors, Anouk, Jean Hars, Eva Faure, Thomas Quintaine, Jean-Yves Chollet, et Sophie Rossi. 2014. « Le réseau Sagir: un outil de vigilance vis-à-vis des agents pathogènes exotiques ». *Bulletin épidémiologique, santé animale et alimentation*, n° 66: 35-39.
- DEFRA/AHS/BEVA. 2015. « Equine quarterly surveillance reports ». http://www.aht.org.uk/cms-display/DEFRA_AHT_BEVA_equine_reports.html.
- de Koeijer, Aline A, Gert Jan Boender, Gonnie Nodelijk, Christoph Staubach, Estelle Meroc, et Armin R W Elbers. 2011. « Quantitative Analysis of Transmission Parameters for Bluetongue Virus Serotype 8 in Western Europe in 2006 ». *Veterinary Research* 42: 53. doi:10.1186/1297-9716-42-53.
- de Koeijer, Aline, Bram Schreuder, et Annemarie Bouma. 2002. « Factors that influence the age distribution of BSE cases: potentials for age targeting in surveillance ». *Livestock Production Science, Safety of Meat and Milk*, 76 (3): 223-33. doi:10.1016/S0301-6226(02)00122-7.
- Del Amo, Javier, Francisco Llorente, Jordi Figuerola, Ramon C Soriguer, Ana M Moreno, Paolo Cordioli, Herbert Weissenböck, et Miguel Angel Jimenez-Clavero. 2014. « Experimental infection of house sparrows (*Passer domesticus*) with West Nile virus isolates of Euro-Mediterranean and North American origins ». *Veterinary Research* 45 (1): 33. doi:10.1186/1297-9716-45-33.
- Del Amo, Javier, Francisco Llorente, Elisa Pérez-Ramírez, Ramón C. Soriguer, Jordi Figuerola, Norbert Nowotny, et Miguel Angel Jiménez-Clavero. 2014. « Experimental infection of house sparrows (*Passer domesticus*) with West Nile virus strains of lineages 1 and 2 ». *Veterinary Microbiology* 172 (3-4): 542-47. doi:10.1016/j.vetmic.2014.06.005.
- de Vos, C.J., C.A. Hoek, et G. Nodelijk. 2012. « Risk of introducing African horse sickness virus into the Netherlands by international equine movements ». *Preventive Veterinary Medicine* 106 (2): 108-22. doi:10.1016/j.prevetmed.2012.01.019.
- de Vos, C.J., Maarten Hoek, Egil Fischer, Aline De Koeijer, et Johan Bremmer. 2011. « Risk Assessment Framework for Emerging Vector-Borne Livestock Diseases ». 11-CVI0168. Netherlands: Central Veterinary Institute of Wageningen UR.

- Dhama, K., R.V.S. Pawaiya, K. Karthik, S. Chakrabort, R. Tiwari, et A.K. Verma. 2014. « Equine encephalosis virus (EEV): A Review ». *Asian Journal of Animal and Veterinary Advances* 9 (2): 123-33. doi:10.3923/ajava.2014.123.133.
- Diekmann, O., J. a. P. Heesterbeek, et J. a. J. Metz. 1990. « On the Definition and the Computation of the Basic Reproduction Ratio R_0 in Models for Infectious Diseases in Heterogeneous Populations ». *Journal of Mathematical Biology* 28 (4): 365-82. doi:10.1007/BF00178324.
- Di Sabatino, Daria, Rossana Bruno, Francesca Sauro, Maria Luisa Danzetta, Francesca Cito, Simona Iannetti, Valeria Narcisi, Fabrizio De Massis, et Paolo Calistri. 2014. « Epidemiology of West Nile Disease in Europe and in the Mediterranean Basin from 2009 to 2013 ». *BioMed Research International* 2014 (ID 907852): 10. doi:10.1155/2014/907852.
- Dobbins, Maureen, Susan Jack, Helen Thomas, et Anita Kothari. 2007. « Public Health Decision-Makers' Informational Needs and Preferences for Receiving Research Evidence ». *Worldviews on Evidence-Based Nursing* 4 (3): 156-63. doi:10.1111/j.1741-6787.2007.00089.x.
- Dobbins, M., K. DeCorby, et T. Twiddy. 2004. « A Knowledge Transfer Strategy for Public Health Decision Makers ». *Worldviews on Evidence-Based Nursing* 1 (2): 120-28. doi:10.1111/j.1741-6787.2004.t01-1-04009.x.
- Doceul, Virginie, Estelle Lara, Corinne Sailleau, Guillaume Belbis, Jennifer Richardson, Emmanuel Breard, Cyril Viarouge, et al. 2013. « Epidemiology, molecular virology and diagnostics of Schmallenberg virus, an emerging orthobunyavirus in Europe ». *Veterinary Research* 44 (1): 31. doi:10.1186/1297-9716-44-31.
- Doherr, M. G., et L. Audigé. 2001. « Monitoring and Surveillance for Rare Health-Related Events: A Review from the Veterinary Perspective ». *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 356 (1411): 1097-1106. doi:10.1098/rstb.2001.0898.
- Dórea, Fernanda C., Beverly J. McEwen, W. Bruce McNab, Crawford W. Revie, et Javier Sanchez. 2013. « Syndromic Surveillance Using Veterinary Laboratory Data: Data Pre-Processing and Algorithm Performance Evaluation ». *Journal of The Royal Society Interface* 10 (83): 20130114. doi:10.1098/rsif.2013.0114.
- Dórea, Fernanda C., C. Anne Muckle, David Kelton, JT. McClure, Beverly J. McEwen, W. Bruce McNab, Javier Sanchez, et Crawford W. Revie. 2013. « Exploratory Analysis of Methods for Automated Classification of Laboratory Test Orders into Syndromic Groups in Veterinary Medicine ». *PLoS ONE* 8 (3): e57334. doi:10.1371/journal.pone.0057334.
- Dórea, Fernanda C., Crawford W. Revie, Beverly J. McEwen, W. Bruce McNab, David Kelton, et Javier Sanchez. s. d. « Retrospective time series analysis of veterinary laboratory data: Preparing a historical baseline for cluster detection in syndromic surveillance ». *Preventive Veterinary Medicine*. doi:10.1016/j.prevetmed.2012.10.010.
- Dórea, Fernanda C., Javier Sanchez, et Crawford W Revie. 2011. « Veterinary syndromic surveillance: Current initiatives and potential for development ». *Preventive veterinary medicine* 101 (1-2): 1-17. doi:10.1016/j.prevetmed.2011.05.004.
- Douglas, K O, A M Kilpatrick, P N Levett, et M C Lavoie. 2007. « A quantitative risk assessment of West Nile virus introduction into Barbados ». *The West Indian medical journal* 56 (5): 394-97.
- Dridi, Maha, Didier Vangeluwe, Sylvie Lecollinet, Thierry van den Berg, et Bénédicte Lambrecht. 2013. « Experimental infection of Carrion crows (*Corvus corone*) with two European West Nile virus (WNV) strains ». *Veterinary Microbiology, Special Issue: One World, One Health, One Virology*, 165 (1-2): 160-66. doi:10.1016/j.vetmic.2012.12.043.
- Drygajlo, Andrzej, Didier Meuwly, et Anil Alexander. 2003. « Statistical methods and Bayesian interpretation of evidence in forensic automatic speaker recognition. » In *Proc. Eurospeech*.

- http://perso.telecom-paristech.fr/~chollet/Biblio/Cours/Biomet/Eurospeech_2003_AD_DM_AA.pdf.
- Dugas, Andrea Freyer, Yu-Hsiang Hsieh, Scott R. Levin, Jesse M. Pines, Darren P. Mareiniss, Amir Mohareb, Charlotte A. Gaydos, Trish M. Perl, et Richard E. Rothman. 2012. « Google Flu Trends: Correlation With Emergency Department Influenza Rates and Crowding Metrics ». *Clinical Infectious Diseases*, janvier, cir883. doi:10.1093/cid/cir883.
- Dugas, Andrea Freyer, Mehdi Jalalpour, Yulia Gel, Scott Levin, Fred Torcaso, Takeru Igusa, et Richard E Rothman. 2013. « Influenza Forecasting with Google Flu Trends ». *PLoS One* 8 (2): e56176. doi:10.1371/journal.pone.0056176.
- Dupuy, Céline, Anne Bronner, Eamon Watson, Linda Wuyckhuise-Sjouke, Martin Reist, Anne Fouillet, Didier Calavas, Pascal Hendrikx, et Jean-Baptiste Perrin. 2013. « Inventory of veterinary syndromic surveillance initiatives in Europe (Triple-S project): Current situation and perspectives ». *Preventive Veterinary Medicine* 111 (3–4): 220-29. doi:10.1016/j.prevetmed.2013.06.005.
- Dupuy, Céline, Eric Morignat, Xavier Maugey, Jean-Luc Vinard, Pascal Hendrikx, Christian Ducrot, Didier Calavas, et Emilie Gay. 2013. « Defining Syndromes Using Cattle Meat Inspection Data for Syndromic Surveillance Purposes: A Statistical Approach with the 2005–2010 Data from Ten French Slaughterhouses ». *BMC Veterinary Research* 9 (1): 88. doi:10.1186/1746-6148-9-88.
- Durand, Benoit, Sylvie Lecollinet, Cécile Beck, Beatriz Martínez-López, Thomas Balenghien, et Véronique Chevalier. 2013. « Identification of Hotspots in the European Union for the Introduction of Four Zoonotic Arboviroses by Live Animal Trade ». *PLoS ONE* 8 (7): e70000. doi:10.1371/journal.pone.0070000.
- ECDC. 2011. *Evidence-Based Methodologies for Public Health: How to Assess the Best Available Evidence When Time Is Limited and There Is Lack of Sound Evidence*. ECDC Technical Report. Stockholm, Sweden: European Centre for Disease Prevention and Control.
- EFSA. 2009. « Scientific Opinion on Epizootic Hemorrhagic Disease ». 7, EFSA Journal, 12 (1418). <http://www.efsa.europa.eu/en/scdocs/doc/1418.pdf>.
- . 2010. « Guidance on a harmonised framework for pest risk assessment and the identification and evaluation of pest risk management options by EFSA ». *EFSA Journal* 8 (2): 1495-1563. doi:10.2903/j.efsa.2010.1495.
- Eidson, M., L. Kramer, W. Stone, Y. Hagiwara, et K. Schmit. 2001. « Dead bird surveillance as an early warning system for West Nile virus. » *Emerging Infectious Diseases* 7 (4): 631-35.
- El Adlouni, Salaheddine, Claudie Beaulieu, Taha Ouarda, Pierre Gosselin, et André Saint-Hilaire. 2007. « Effects of climate on West Nile Virus transmission risk used for public health decision-making in Quebec ». *International Journal of Health Geographics* 6 (1): 1-7. doi:10.1186/1476-072X-6-40.
- Fan, Yunzhou, Ying Wang, Hongbo Jiang, Wenwen Yang, Miao Yu, Weirong Yan, Vinod K. Diwan, et al. 2014. « Evaluation of Outbreak Detection Performance Using Multi-Stream Syndromic Surveillance for Influenza-Like Illness in Rural Hubei Province, China: A Temporal Simulation Model Based on Healthcare-Seeking Behaviors ». *PLoS ONE* 9 (11): e112255. doi:10.1371/journal.pone.0112255.
- FAO. 2014. « Risk-based surveillance - A manual for veterinarians on the design and analysis of surveillance for demonstration of freedom from disease ». Manual 17. Animal Production and Health. Rome, Italy.
- FAO-OIE-WHO Collaboration. 2006. « Global Early Warning and Response for Major Animal Diseases, including Zoonoses (GLEWS) ». <http://www.oie.int/doc/ged/D11304.PDF>.

- . 2013. « GLEWS, The Joint FAO–OIE–WHO Global Early Warning System for health threats and emerging risks at the human–animal–ecosystems interface ». http://www.glews.net/wp-content/uploads/2013/12/04_GLEWSConcept-20-11.pdf.
- Farrington, C. P., N. J. Andrews, A. D. Beale, et M. A. Catchpole. 1996. « A Statistical Algorithm for the Early Detection of Outbreaks of Infectious Disease ». *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 159 (3): 547-63. doi:10.2307/2983331.
- Faverjon, Céline, Mats Gunnar Andersson, Jackie Tapprest, Carole Sala, Anouk Decors, Pierre Triz, Orsolya Kutasi, Alain Sandoz, et Agnès Leblond. 2015. « Evaluation of a multivariate syndromic surveillance system for West Nile virus », Manuscript submitted, .
- Faverjon, C., A. Leblond, P. Hendriks, T. Balenghien, C. J. de Vos, E.A.J. Fischer, et A.A. de Koeijer. 2015. « A Spatiotemporal Model to Assess the Introduction Risk of African Horse Sickness by Import of Animals and Vectors in France ». *BMC Veterinary Research* 11 (1). doi:10.1186/s12917-015-0435-4.
- Fischer, Egil AJ, Gert-Jan Boender, Gonnie Nodelijk, Aline A. de Koeijer, et H. J. Van Roermund. 2013. « The transmission potential of Rift Valley fever virus among livestock in the Netherlands: a modelling study ». *Vet Res* 44 (1): 58. doi:10.1186/1297-9716-44-58.
- Foreman, L. A., C. Champod, I. W. Evett, J. A. Lambert, et S. Pope. 2003. « Interpreting DNA Evidence: A Review ». *International Statistical Review* 71 (3): 473-95. doi:10.1111/j.1751-5823.2003.tb00207.x.
- Frisén, Marianne, Eva Andersson, et Linus Schiöler. 2010. « Evaluation of multivariate surveillance ». *Journal of Applied Statistics* 37 (12): 2089-2100. doi:10.1080/02664760903222208.
- Gage, Kenneth L., Thomas R. Burkot, Rebecca J. Eisen, et Edward B. Hayes. 2008. « Climate and Vectorborne Diseases ». *American Journal of Preventive Medicine*, Theme Issue: Climate Change and the Health of the Public, 35 (5): 436-50. doi:10.1016/j.amepre.2008.08.030.
- Gardner, Everette S. 1985. « Exponential Smoothing: The State of the Art ». *Journal of Forecasting* 4 (1): 1-28. doi:10.1002/for.3980040103.
- Gesualdo, Francesco, Giovanni Stilo, Eleonora Agricola, Michaela V Gonfiantini, Elisabetta Pandolfi, Paola Velardi, et Alberto E Tozzi. 2013. « Influenza-like Illness Surveillance on Twitter through Automated Learning of Naïve Language ». *PloS One* 8 (12): e82489. doi:10.1371/journal.pone.0082489.
- Ginsberg, Jeremy, Matthew H. Mohebbi, Rajan S. Patel, Lynnette Brammer, Mark S. Smolinski, et Larry Brilliant. 2009. « Detecting Influenza Epidemics Using Search Engine Query Data ». *Nature* 457 (7232): 1012-14. doi:10.1038/nature07634.
- Giovannini, Armando, Stuart MacDiarmid, Paolo Calistri, Annamaria Conte, Lara Savini, Donatella Nannini, et Steve Weber. 2004. « The Use of Risk Assessment to Decide the Control Strategy for Bluetongue in Italian Ruminant Populations ». *Risk Analysis* 24 (6): 1737-53. doi:10.1111/j.0272-4332.2004.00563.x.
- Gittelsohn, Simone N. 2013. « Evolving from Inferences to Decisions in the Interpretation of Scientific Evidence ». Thesis, Lausanne: Université de Lausanne, Faculté de Droit et des Sciences Criminelles Ecole des Sciences Criminelles. http://www.unil.ch/esc/files/live/sites/esc/files/shared/These_Gittelsohn.pdf.
- Gosselin, Pierre, Germain Lebel, Sonia Rivest, et Monique Douville-Fradet. 2005. « The Integrated System for Public Health Monitoring of West Nile Virus (ISPHM-WNV): a real-time GIS for surveillance and decision-making ». *International Journal of Health Geographics* 4 (1): 1-12. doi:10.1186/1476-072X-4-21.

- Gratz, N. G. 2004. « Critical Review of the Vector Status of *Aedes Albopictus* ». *Medical and Veterinary Entomology* 18 (3): 215-27. doi:10.1111/j.0269-283X.2004.00513.x.
- Greene, Sharon K., Jie Huang, Allyson M. Abrams, Debra Gilliss, Mary Reed, Richard Platt, Susan S. Huang, et Martin Kulldorff. 2012. « Gastrointestinal Disease Outbreak Detection Using Multiple Data Streams from Electronic Medical Records ». *Foodborne Pathogens and Disease* 9 (5): 431-41. doi:10.1089/fpd.2011.1036.
- Gubbins, Simon, Simon Carpenter, Matthew Baylis, James L.N Wood, et Philip S Mellor. 2008. « Assessing the risk of bluetongue to UK livestock: uncertainty and sensitivity analyses of a temperature-dependent model for the basic reproduction number ». *Journal of the Royal Society Interface* 5 (20): 363-71. doi:10.1098/rsif.2007.1110.
- Guis, Helene, Cyril Caminade, Carlos Calvete, Andrew P. Morse, Annelise Tran, et Matthew Baylis. 2012. « Modelling the Effects of Past and Future Climate on the Risk of Bluetongue Emergence in Europe ». *Journal of The Royal Society Interface* 9 (67): 339-50. doi:10.1098/rsif.2011.0255.
- Gustafson, L., K. Klotins, S. Tomlinson, G. Karreman, A. Cameron, B. Wagner, M. Remmenga, N. Bruneau, et A. Scott. 2010. « Combining Surveillance and Expert Evidence of Viral Hemorrhagic Septicemia Freedom: A Decision Science Approach ». *Preventive Veterinary Medicine* 94 (1-2): 140-53. doi:10.1016/j.prevetmed.2009.11.021.
- Hadorn, Daniela C., Sabina S. Haracic, et Katharina DC Stärk. 2008. « Comparative Assessment of Passive Surveillance in Disease-Free and Endemic Situation: Example of *Brucella Melitensis* Surveillance in Switzerland and in Bosnia and Herzegovina ». *BMC Veterinary Research* 4 (1): 52. doi:10.1186/1746-6148-4-52.
- Hanley, J. A., et B. J. McNeil. 1982. « The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve ». *Radiology* 143 (1): 29-36. doi:10.1148/radiology.143.1.7063747.
- Hartemink, N.A., B.V. Purse, R. Meiswinkel, H.E. Brown, A. de Koeijer, A.R.W. Elbers, G.-J. Boender, D.J. Rogers, et J.A.P. Heesterbeek. 2009. « Mapping the basic reproduction number (R0) for vector-borne diseases: A case study on bluetongue virus ». *Epidemics* 1 (3): 153-61. doi:10.1016/j.epidem.2009.05.004.
- Hartnett, Emma, Amie Adkin, Miles Seaman, John Cooper, Eamon Watson, Helen Coburn, Tracey England, Christophen Marooney, Anthony Cox, et Mavion Wooldridge. 2007. « A Quantitative Assessment of the Risks from Illegally Imported Meat Contaminated with Foot and Mouth Disease Virus to Great Britain ». *Risk Analysis* 27 (1): 187-202. doi:10.1111/j.1539-6924.2006.00869.x.
- Havelaar, Arie H., Floor van Rosse, Catalin Bucura, Milou A. Toetenel, Juanita A. Haagsma, Dorota Kurowicka, J. (Hans) A. P. Heesterbeek, et al. 2010. « Prioritizing Emerging Zoonoses in The Netherlands ». *PLoS ONE* 5 (11): e13965. doi:10.1371/journal.pone.0013965.
- Henderson, Donald A., Thomas V. Inglesby, Tara O'Toole, Annie Fine, et Marcelle Layton. 2001. « Lessons from the West Nile Viral Encephalitis Outbreak in New York City, 1999: Implications for Bioterrorism Preparedness ». *Clinical Infectious Diseases* 32 (2): 277-82. doi:10.1086/318469.
- Hendrickx, Guy, Marius Gilbert, Christoph Staubach, Armin Elbers, Koen Mintiens, Guillaume Gerbier, et Els Ducheyne. 2008. « A wind density model to quantify the airborne spread of *Culicoides* species during north-western Europe bluetongue epidemic, 2006 ». *Preventive Veterinary Medicine* 87 (1-2): 162-81. doi:10.1016/j.prevetmed.2008.06.009.
- Hernández-Triana, Luis M., Claire L. Jeffries, Karen L. Mansfield, George Carnell, Anthony R. Fooks, et Nicholas Johnson. 2014. « Emergence of West Nile Virus Lineage 2 in Europe: A Review on the Introduction and Spread of a Mosquito-Borne Disease ». *Frontiers in Public Health* 2 (271). doi:10.3389/fpubh.2014.00271.

- Hiller, Katherine M, Lisa Stoneking, Alice Min, et Suzanne Michelle Rhodes. 2013. « Syndromic Surveillance for Influenza in the Emergency Department-A Systematic Review ». *PloS One* 8 (9): e73832. doi:10.1371/journal.pone.0073832.
- Hoar, Bruce, Tim Carpenter, Randall Singer, et Ian Gardner. 2004. « Probability of introduction of exotic strains of bluetongue virus into the US and into California through importation of infected cattle ». *Preventive Veterinary Medicine* 66 (1-4): 79-91. doi:10.1016/j.prevetmed.2004.08.006.
- Hoffman, F. Owen, et Jana S. Hammonds. 1994. « Propagation of Uncertainty in Risk Assessments: The Need to Distinguish Between Uncertainty Due to Lack of Knowledge and Uncertainty Due to Variability ». *Risk Analysis* 14 (5): 707-12. doi:10.1111/j.1539-6924.1994.tb00281.x.
- Höle, Michael, Michaela Paul, et Leonhard Held. 2007. « Statistical approaches to the surveillance of infectious diseases for veterinary public health ». Technical report 14. Munich, Germany: Department of Statistics, University of Munich. <http://core.ac.uk/download/pdf/12163041.pdf>.
- Hunter, J.S. 1986. « The exponentially weighted moving average », *Journal of Quality Technology*, 18 (4): 203-210.
- Hutwagner, L. C., W. W. Thompson, G. M. Seeman, et T. Treadwell. 2005. « A Simulation Model for Assessing Aberration Detection Methods Used in Public Health Surveillance for Systems with Limited Baselines ». *Statistics in Medicine* 24 (4): 543-50. doi:10.1002/sim.2034.
- IFCE - les Haras nationaux. 2011. *Panorama économique de la filière équine*. IFCE - les Haras nationaux.
- Ivanov, Oleg, Michael M. Wagner, Wendy W. Chapman, et Robert T. Olszewski. 2002. « Accuracy of three classifiers of acute gastrointestinal syndrome for syndromic surveillance. » In *Proceedings of AMIA annual Symposium*, 345-49. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2244375/>.
- Jacobs, Ja, E Jones, Ba Gabella, B Spring, et Rc Brownson. 2012. « Tools for Implementing an Evidence-Based Approach in Public Health Practice ». *Preventing Chronic Disease* 9 (110324). doi:10.5888/pcd9.110324.
- Johnson, Glen D., Millicent Eidson, Kathryn Schmit, April Ellis, et Martin Kulldorff. 2006. « Geographic Prediction of Human Onset of West Nile Virus Using Dead Crow Clusters: An Evaluation of Year 2002 Data in New York State ». *American Journal of Epidemiology* 163 (2): 171-80. doi:10.1093/aje/kwj023.
- Jones, Kate E., Nikkita G. Patel, Marc A. Levy, Adam Storeygard, Deborah Balk, John L. Gittleman, et Peter Daszak. 2008. « Global Trends in Emerging Infectious Diseases ». *Nature* 451 (7181): 990-93. doi:10.1038/nature06536.
- Jones, Robert H., et N. M. Foster. 1971. « The effect of repeated blood meals for Bluetongue on the infection rate of *Culicoides variipennis* ». *Journal of Medical Entomology* 8 (5): 499-501. doi:http://dx.doi.org/10.1093/jmedent/8.5.499.
- Kalekar, Prajakta S. 2004. « Time series forecasting using holt-winters exponential smoothing ». *Kanwal Rekhi School of Information Technology* 4329008: 1-13.
- Katz, Rebecca, Larissa May, Julia Baker, et Elisa Test. 2011. « Redefining syndromic surveillance ». *Journal of Epidemiology and Global Health* 1 (1): 21-31. doi:10.1016/j.jegh.2011.06.003.
- Kazeem, M.M., N. Rufai, E.A. Ogunsan, L.H. Lombin, L.U. Enurah, et O. Owolodun. 2008. « Clinicopathological Features Associated with the Outbreak of African Horse Sickness in Lagos, Nigeria ». *Journal of Equine Veterinary Science* 28 (10): 594-97. doi:10.1016/j.jevs.2008.08.002.
- Kilpatrick, A.Marm, Yekaterina Gluzberg, Jeff Burgett, et Peter Daszak. 2004. « Quantitative Risk Assessment of the Pathways by Which West Nile Virus Could Reach Hawaii ». *EcoHealth* 1 (2): 205-9. doi:10.1007/s10393-004-0086-6.

- Kosmider, R., L. Kelly, S. Evans, et G. Gettinby. 2006. « A statistical system for detecting Salmonella outbreaks in British livestock ». *Epidemiology and Infection* 134 (5): 952-60. doi:10.1017/S0950268806005887.
- Kramer, Laura D., et Kristen A. Bernard. 2001. « West Nile Virus Infection in Birds and Mammals ». *Annals of the New York Academy of Sciences* 951 (1): 84-93. doi:10.1111/j.1749-6632.2001.tb02687.x.
- Krause, Gérard. 2008. « How Can Infectious Diseases Be Prioritized in Public Health? » *EMBO Reports* 9 (S1): S22-27. doi:10.1038/embo.2008.76.
- Kulasekera, V. L., L. Kramer, R. S. Nasci, F. Mostashari, B. Cherry, S. C. Trock, C. Glaser, et J. R. Miller. 2001. « West Nile virus infection in mosquitoes, birds, horses, and humans, Staten Island, New York, 2000. » *Emerging Infectious Diseases* 7 (4): 722-25.
- Kulldorff, Martin, Farzad Mostashari, Luiz Duczmal, Katherine Yih, Ken Kleinman, et Richard Platt. 2007. « Multivariate Scan Statistics for Disease Surveillance ». *Statistics in Medicine: Special Issue: Tenth Biennial CDC and ATSDR Symposium on Statistical Methods* 26 (8): 1824-33. doi:10.1002/sim.2818.
- Kutasi, O., T. Bakonyi, S. Lecollinet, I. Biksi, E. Ferenczi, C. Bahuon, S. Sardi, S. Zientara, et O. Szenci. 2011. « Equine Encephalomyelitis Outbreak Caused by a Genetic Lineage 2 West Nile Virus in Hungary: Lineage 2 West Nile Virus Encephalomyelitis ». *Journal of Veterinary Internal Medicine* 25 (3): 586-91. doi:10.1111/j.1939-1676.2011.0715.x.
- Latham, J., L. Murajda, F. Forland, et A. Jansen. 2013. « Capacities, practices and perceptions of evidence-based public health in Europe ». *Euro Surveillance* 18 (10): 13.
- Lau, Eric H. Y., Benjamin J. Cowling, Lai-Ming Ho, et Gabriel M. Leung. 2008. « Optimizing Use of Multistream Influenza Sentinel Surveillance Data ». *Emerging Infectious Diseases* 14 (7): 1154-57. doi:10.3201/eid1407.080060.
- Leadon, D. P., et C. P. Herholz. 2009. « Globalisation of trade and the spread of infectious disease ». *Equine vet. Educ. Manual* 8. doi:10.2746/095777309X463640.
- Leblond, A., P. Hendrikx, et P. Sabatier. 2007. « West Nile Virus Outbreak Detection Using Syndromic Monitoring in Horses ». *Vector-Borne and Zoonotic Diseases* 7 (3): 403-10. doi:10.1089/vbz.2006.0593.
- Lebrun, Jacky. 2010. « Les enjeux et les perspectives de la filière équine en France ». Rapport public 9. Conseil économique, social et environnemental. Section de l'agriculture et de l'alimentation. <http://www.ladocumentationfrancaise.fr/rapports-publics/104000387/index.shtml>.
- Léon, A, G. Fortier, C. Fortier, F. Freymuth, Jackie Tapprest, R Leclercq, et S. Pronost. 2008. « Detection of equine herpesviruses in aborted fetuses by consensus PCR ». *Veterinary Microbiology* 126 (1): 20. doi:10.1016/j.vetmic.2007.06.019.
- Levin-Rector, Alison, Elisha L. Wilson, Annie D. Fine, et Sharon K. Greene. 2015. « Refining Historical Limits Method to Improve Disease Cluster Detection, New York City, New York, USA ». *Emerging Infectious Diseases* 21 (2): 265-72. doi:10.3201/eid2102.140098.
- Liljenstolpe, Carolina. 2009. « Horses in Europe ». Sweden: Swedish University of Agricultural Sciences. <http://www.wbfsh.org/files/EU%20Equus%202009.pdf>.
- Lindgren, Elisabet, Yvonne Andersson, Jonathan E. Suk, Bertrand Sudre, et Jan C. Semenza. 2012. « Monitoring EU emerging infectious disease risk due to climate change ». *Science* 336 (6080): 418-19. doi:10.1126/science.1215735.
- Lo Iacono, G., C. A. Robin, J. R. Newton, S. Gubbins, et J. L. N. Wood. 2013. « Where are the horses? With the sheep or cows? Uncertain host location, vector-feeding preferences and the risk of African

- horse sickness transmission in Great Britain ». *Journal of The Royal Society Interface* 10 (83): 20130194-20130194. doi:10.1098/rsif.2013.0194.
- Lomas, Jonathan. 1997. « Research and Evidence-based Decision Making ». *Australian and New Zealand Journal of Public Health* 21 (5): 439-41. doi:10.1111/j.1467-842X.1997.tb01730.x.
- Lord, C. C., G. J. Venter, P. S. Mellor, J. T. Paweska, et M. E. J. Woolhouse. 2002. « Transmission patterns of African horse sickness and equine encephalosis viruses in South African donkeys ». *Epidemiology and infection* 128 (02): 265-75. doi:10.1017 / S0950268801006471.
- Lord, C. C., M. E. J. Woolhouse, J. A. P. Heesterbeek, et P. S. Mellor. 1996. « Vector-borne diseases and the basic reproduction number: a case study of African horse sickness ». *Medical and veterinary entomology* 10 (1): 19-28. doi:10.1111/j.1365-2915.1996.tb00077.x.
- Lotze, Thomas, Sean Murphy, et Galit Shmueli. 2008. « Implementation and comparison of preprocessing methods for biosurveillance data ». *Advances in Disease Surveillance* 6 (1): 1-20.
- Lunn, D.P., N. Davis-Poynter, M.J.B.F. Flaminio, D.W. Horohov, K. Osterrieder, N. Pusterla, et H.G.G. Townsend. 2009. « Equine Herpesvirus-1 Consensus Statement ». *Journal of Veterinary Internal Medicine* 23 (3): 450-61. doi:10.1111/j.1939-1676.2009.0304.x.
- Macdonald, G. 1955. « The Measurement of Malaria Transmission ». In *Proceedings of the Royal Society of Medicine*, 48(4):295-302. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1918775/>.
- MacGregor, J. F., et T. Kourti. 1995. « Statistical process control of multivariate processes ». *Control Engineering Practice* 3 (3): 403-14. doi:10.1016/0967-0661(95)00014-L.
- Maciejewski, Ross, Benjamin Tyner, Yun Jang, Cheng Zheng, Rimma Nehme, David Ebert, William Cleveland, Mourad Ouzzani, Shaun Grannis, et Lawrence Glickman. 2007. « LAHVA: Linked Animal-Human Health Visual Analytics ». In *Proceedings of the IEEE Symposium on Visual Analytics Science and Technology*, 27-34. Sacramento, California USA: William Ribarsky, John Dill. doi:10.1109/VAST.2007.4388993.
- MacLachlan, N. James, et Alan J. Guthrie. 2010. « Re-emergence of bluetongue, African horse sickness, and other Orbivirus diseases ». *Veterinary Research* 41 (6). doi:10.1051/vetres/2010007.
- Madouasse, Aurélien, Alexis Marceau, Anne Lehébel, Henriëtte Brouwer-Middelesch, Gerdien van Schaik, Yves Van der Stede, et Christine Fourichon. 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. doi:10.1371/journal.pone.0073726.
- Mandl, Kenneth D., J. Marc Overhage, Michael M. Wagner, William B. Lober, Paola Sebastiani, Farzad Mostashari, Julie A. Pavlin, et al. 2004. « Implementing Syndromic Surveillance: A Practical Guide Informed by the Early Experience ». *Journal of the American Medical Informatics Association : JAMIA* 11 (2): 141-50. doi:10.1197/jamia.M1356.
- Mandl, Kenneth D, B Reis, et C Cassa. 2004. « Measuring outbreak-detection performance by using controlled feature set simulations ». *MMWR, Morbidity and Mortality Weekly Report* 53 Suppl: 130-36.
- Marsh, Mercer, Kroll, Guy Carpenter, et Olivier Wyman. 2008. « The economic and social impact of emerging infectious diseases: Mitigation through detection, research, and response ». London: Marsh Inc. http://www.healthcare.philips.com/main/shared/assets/documents/bioshield/ecoandsocialimpactofemerginginfectiousdisease_111208.pdf.
- Martin, P. a. J., A. R. Cameron, K. Barfod, E. S. G. Sergeant, et M. Greiner. 2007. « Demonstrating Freedom from Disease Using Multiple Complex Data Sources 2: Case Study--Classical Swine Fever in

- Denmark ». *Preventive Veterinary Medicine* 79 (2-4): 98-115.
doi:10.1016/j.prevetmed.2006.09.007.
- Martin, P. A. J., A. R. Cameron, et M. Greiner. 2007. « Demonstrating freedom from disease using multiple complex data sources: 1: A new methodology based on scenario trees ». *Preventive Veterinary Medicine* 79 (2-4): 71-97. doi:10.1016/j.prevetmed.2006.09.008.
- Massad, Eduardo, Francisco Antonio Bezerra Coutinho, Luis Fernandez Lopez, et Daniel Rodrigues da Silva. 2011. « Modeling the Impact of Global Warming on Vector-Borne Infections ». *Physics of Life Reviews* 8 (2): 169-99. doi:10.1016/j.plrev.2011.01.001.
- McKenzie, Joanna, Helen Simpson, et Ian Langstaff. 2007. « Development of methodology to prioritise wildlife pathogens for surveillance ». *Preventive Veterinary Medicine* 81 (1-3): 194-210. doi:10.1016/j.prevetmed.2007.04.003.
- Mellor, Philip. Scott, et Christopher Hamblin. 2004. « African horse sickness ». *Veterinary research* 35 (4): 445-66. doi:10.1051/vetres:2004021.
- Mildenberg, Z., D. Westcott, M. Bellaiche, A. Dastjerdi, F. Steinbach, et T. Drew. 2009. « Equine Encephalosis Virus in Israel ». *Transboundary and Emerging Diseases* 56 (8): 291-291. doi:10.1111/j.1865-1682.2009.01087_1.x.
- Morrison, Geoffrey Stewart. 2012. « The Likelihood-Ratio Framework and Forensic Evidence in Court: A Response to *R v T* ». *The International Journal of Evidence & Proof* 16 (1): 1-29. doi:10.1350/ijep.2012.16.1.390.
- Mostashari, Farzad, Martin Kulldorff, Jessica J. Hartman, James R. Miller, et Varuni Kulasekera. 2003. « Dead Bird Clusters as an Early Warning System for West Nile Virus Activity ». *Emerging Infectious Diseases* 9 (6): 641-46. doi:10.3201/eid0906.020794.
- Murgue, B., S. Murri, S. Zientara, B. Durand, J. P. Durand, et H. Zeller. 2001. « West Nile outbreak in horses in southern France, 2000: the return after 35 years. » *Emerging Infectious Diseases* 7 (4): 692-96. doi:10.3201/eid0704.010417.
- Murray, Gardner, Suzanne Munstermann, et Kenneth Lam. 2013. « benefits and challenges posed by the worldwide expansion of equestrian events. New standards for the population of competition horses and equine disease free zones (EDFZ) in countries ». In *81st general session, world assembly of the OIE*. Paris.
http://www.oie.int/fileadmin/Home/eng/Publications_%26_Documentation/docs/pdf/TT/2013_A_81SG_10.pdf.
- Muscatello, David J., Patricia M. Morton, Ingrid Evans, et Robin Gilmour. 2008. « Prospective Surveillance of Excess Mortality due to Influenza in New South Wales: Feasibility and Statistical Approach ». *Communicable Diseases Intelligence Quarterly Report* 32 (4): 435-42.
- Napp, S, I García-Bocanegra, N Pagès, A Allepuz, A Alba, et J Casal. 2012. « Assessment of the risk of a bluetongue outbreak in Europe caused by *Culicoides* midges introduced through intracontinental transport and trade networks ». *Medical and veterinary entomology* 27 (1): 19-28. doi:10.1111/j.1365-2915.2012.01016.x.
- Ninio, Camille, Denis Augot, Jean-Claude Delecolle, Barbara Dufour, et Jerome Depaquit. 2011. « Contribution to the Knowledge of *Culicoides* (Diptera: Ceratopogonidae) Host Preferences in France ». *Parasitology Research* 108 (3): 657-63. doi:10.1007/s00436-010-2110-9.
- Nordin, James D., Michael J. Goodman, Martin Kulldorff, Debra P. Ritzwoller, Allyson M. Abrams, Ken Kleinman, Mary Jeanne Levitt, James Donahue, et Richard Platt. 2005. « Simulated Anthrax Attacks and Syndromic Surveillance ». *Emerging Infectious Diseases* 11 (9): 1394-98. doi:10.3201/eid1109.050223.

- O'Brien, SJ, et P Christie. 1997. « Do CuSums have a role in routine communicable disease surveillance? » *Public Health* 111 (4): 255-58. doi:10.1038/sj.ph.1900353.
- Oidtman, Birgit, Edmund Peeler, Trude Lyngstad, Edgar Brun, Britt Bang Jensen, et Katharina D. C. Stärk. 2013. « Risk-based methods for fish and terrestrial animal disease surveillance ». *Preventive Veterinary Medicine* 112 (1-2): 13-26. doi:10.1016/j.prevetmed.2013.07.008.
- OIE. 2010. « Import Risk Analysis ». In *International Animal Health Code*. http://web.oie.int/eng/normes/mcode/en_chapitre_1.2.1.pdf.
- . 2014. « Import Risk Analysis ». In *International Animal Health Code*. OIE. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.431.2817&rep=rep1&type=pdf>.
- Olson, Karen L., Marco Bonetti, Marcello Pagano, et Kenneth D. Mandl. 2005. « Real time spatial cluster detection using interpoint distances among precise patient locations ». *BMC medical informatics and decision making* 5 (1): 19. doi:10.1186/1472-6947-5-19.
- Ozdenerol, Esra, Gregory N. Taff, et Cem Akkus. 2013. « Exploring the Spatio-Temporal Dynamics of Reservoir Hosts, Vectors, and Human Hosts of West Nile Virus: A Review of the Recent Literature ». *International Journal of Environmental Research and Public Health* 10 (11): 5399-5432. doi:10.3390/ijerph10115399.
- Ozonoff, Al, Caroline Jeffery, Justin Manjourides, Laura Forsberg White, et Marcello Pagano. 2007. « Effect of Spatial Resolution on Cluster Detection: A Simulation Study ». *International Journal of Health Geographics* 6 (1): 52. doi:10.1186/1476-072X-6-52.
- Palisade Corporation. 2013. *@Risk Version 6.1. Risk, Analysis and Simulation Add-In for Microsoft Excel* (version 6.1). Risk, Analysis and Simulation Add-In for Microsoft Excel. Palisade Corporation.
- Paweska, J. T., et G. J. Venter. 2004. « Vector Competence of Culicoides Species and the Seroprevalence of Homologous Neutralizing Antibody in Horses for Six Serotypes of Equine Encephalosis Virus (EEV) in South Africa ». *Medical and Veterinary Entomology* 18 (4): 398-407. doi:10.1111/j.0269-283X.2004.00524.x.
- Penzhorn, B. L. 1985. « Reproductive characteristics of a free-ranging population of Cape mountain zebra (*Equus zebra zebra*) ». *Journal of reproduction and fertility* 73 (1): 51-57. doi:10.1530/jrf.0.0730051.
- Pérez-Ramírez, Elisa, Francisco Llorente, et Miguel Ángel Jiménez-Clavero. 2014. « Experimental Infections of Wild Birds with West Nile Virus ». *Viruses* 6 (2): 752-81. doi:10.3390/v6020752.
- Perrin, Jean-Baptiste, Christian Ducrot, Jean-Luc Vinard, Eric Morignat, Didier Calavas, et Pascal Hendrikx. 2012. « Assessment of the utility of routinely collected cattle census and disposal data for syndromic surveillance ». *Preventive Veterinary Medicine* 105 (3): 244-52. doi:10.1016/j.prevetmed.2011.12.015.
- Perrin, Jean-Baptiste, Christian Ducrot, Jean-Luc Vinard, Eric Morignat, André Gauffier, Didier Calavas, et Pascal Hendrikx. 2010. « Using the National Cattle Register to estimate the excess mortality during an epidemic: Application to an outbreak of Bluetongue serotype 8 ». *Epidemics* 2 (4): 207-14. doi:10.1016/j.epidem.2010.10.002.
- Ponçon, Nicolas, Annelise Tran, Céline Toty, Adrian Luty, et Didier Fontenille. 2008. « A Quantitative Risk Assessment Approach for Mosquito-Borne Diseases: Malaria Re-Emergence in Southern France ». *Malaria Journal* 7 (1): 147. doi:10.1186/1475-2875-7-147.
- Portas, M., F. S. Boinas, J. Oliveira E Sousa, et P. Rawlings. 1999. « African horse sickness in Portugal: a successful eradication programme. » *Epidemiology and Infection* 123 (2): 337-46.

- Purse, Bethan V., Philip S. Mellor, David J. Rogers, Alan R. Samuel, Peter P. C. Mertens, et Matthew Baylis. 2005. « Climate Change and the Recent Emergence of Bluetongue in Europe ». *Nature Reviews Microbiology* 3 (2): 171-81. doi:10.1038/nrmicro1090.
- Reisen, William K. 2010. « Landscape Epidemiology of Vector-Borne Diseases ». *Annual Review of Entomology* 55 (1): 461-83. doi:10.1146/annurev-ento-112408-085419.
- Robin, C. A., C. E. Wylie, J. L. N. Wood, et J. R. Newton. 2011. « Making Use of Equine Population Demography for Disease Control Purposes: Preliminary Observations on the Difficulties of Counting and Locating Horses in Great Britain ». *Equine Veterinary Journal* 43 (3): 372-75. doi:10.1111/j.2042-3306.2010.00186.x.
- Rockx, Barry, Liselotte van Asten, Cees van den Wijngaard, Gert-Jan Godeke, Lutz Goehring, Harry Vennema, Harrie van der Avoort, Wilfrid van Pelt, et Marion Koopmans. 2006. « Syndromic surveillance in the Netherlands for the early detection of West Nile virus epidemics ». *Vector borne and zoonotic diseases* 6 (2): 161-69. doi:10.1089/vbz.2006.6.161.
- Rodriguez, M., H. Hooghuis, et Ma. Castaño. 1992. « African horse sickness in Spain ». *Veterinary Microbiology* 33 (1-4): 129-42. doi:10.1016/0378-1135(92)90041-Q.
- Rodriguez, M., J.L. Ladero, M. Castaño, et H. Hooghuis. 1992. « African horse sickness in Spain: Epizootiological and regulatory considerations ». *Journal of Equine Veterinary Science* 12 (6): 395-400. doi:10.1016/S0737-0806(06)81369-9.
- Rosà, Roberto, Giovanni Marini, Luca Bolzoni, Markus Neteler, Markus Metz, Luca Delucchi, Elizabeth A. Chadwick, et al. 2014. « Early warning of West Nile virus mosquito vector: climate and land use models successfully explain phenology and abundance of Culex pipiens mosquitoes in north-western Italy ». *Parasit. Vectors* 7 (269). doi:10.1186/1756-3305-7-269.
- Roure, Josep, Artur Dubrawski, et Jeff Schneider. 2007. « A study into detection of bio-events in multiple streams of surveillance data ». *Intelligence and Security Informatics: Biosurveillance* 4506: 124-33. doi:10.1007/978-3-540-72608-1_12.
- Rutherford, George W., William McFarland, Hilary Spindler, Karen White, Sadhna V. Patel, John Aberle-Grasse, Keith Sabin, et al. 2010. « Public Health Triangulation: Approach and Application to Synthesizing Data to Understand National and Local HIV Epidemics ». *BMC Public Health* 10 (1): 447. doi:10.1186/1471-2458-10-447.
- Ruzante, Juliana Martins, Valerie J. Davidson, Julie Caswell, Aamir Fazil, John A. L. Cranfield, Spencer J. Henson, Sven M. Anders, Claudia Schmidt, et Jeffrey M. Farber. 2010. « A Multifactorial Risk Prioritization Framework for Foodborne Pathogens ». *Risk Analysis* 30 (5): 724-42. doi:10.1111/j.1539-6924.2009.01278.x.
- Rychetnik, Lucie, Penelope Hawe, Elizabeth Waters, Alexandra Barratt, et Michael Frommer. 2004. « A Glossary for Evidence Based Public Health ». *Journal of Epidemiology and Community Health* 58 (7): 538-45. doi:10.1136/jech.2003.011585.
- Saegerman, C., A. Alba-Casals, I. García-Bocanegra, F. Dal Pozzo, et G. van Galen. 2014. « Clinical Sentinel Surveillance of Equine West Nile Fever, Spain ». *Transboundary and Emerging Diseases*. doi:10.1111/tbed.12243.
- Salman, M.D. 2003. « Animal Disease Surveillance and Survey Systems ». In *Animal Disease Surveillance and Survey Systems*, édité par M. D. Salman, 3-13. Iowa State Press.
- Scheffer, Elisabeth G., Gert J. Venter, Karien Labuschagne, Patrick C. Page, Bradley A. Mullens, N. James MacLachlan, Nikolaus Osterrieder, et Alan J. Guthrie. 2012. « Comparison of two trapping methods for Culicoides biting midges and determination of African horse sickness virus prevalence in midge populations at Onderstepoort, South Africa ». *Veterinary Parasitology* 185 (2-4): 265-73. doi:10.1016/j.vetpar.2011.09.037.

- Schiöler, Linus, et Marianne Frisé. 2011. « Multivariate outbreak detection ». *Journal of Applied Statistics* 39 (2): 223-42. doi:10.1080/02664763.2011.584522.
- Serfling, Robert E. 1963. « Methods for current statistical analysis of excess pneumonia-influenza deaths ». *Public Health Reports* 78 (6): 494-506.
- Shaffer, Loren E. 2007. « Using pre-diagnostic data from veterinary laboratories to detect disease outbreaks in companion animals ». Thesis, Ohio State University. http://etd.ohiolink.edu/view.cgi?acc_num=osu1176376010.
- Shephard, Richard William. 2006. « The development of a syndromic surveillance system for the extensive beef cattle producing regions of Australia ». Thesis, University of Sydney. Faculty of Veterinary Science. <http://ses.library.usyd.edu.au:80/handle/2123/2210>.
- Shewhart, Walter Andrew. 1931. *Economic Control of Quality of Manufactured Product*. Reprinted, republished. ASQ Quality Press.
- Shmueli, Galit, et Howard Burkom. 2010. « Statistical Challenges Facing Early Outbreak Detection in Biosurveillance ». *Technometrics* 52 (1): 39-51. doi:10.1198/TECH.2010.06134.
- Shuai, Jiangping, Peter Buck, Paul Sockett, Jeff Aramini, et Frank Pollari. 2006. « A GIS-driven integrated real-time surveillance pilot system for national West Nile virus dead bird surveillance in Canada ». *International Journal of Health Geographics* 5 (1): 1-14. doi:10.1186/1476-072X-5-17.
- Signorini, Alessio, Alberto Maria Segre, et Philip M. Polgreen. 2011. « The Use of Twitter to Track Levels of Disease Activity and Public Concern in the U.S. during the Influenza A H1N1 Pandemic ». *PLoS ONE* 6 (5): e19467. doi:10.1371/journal.pone.0019467.
- Sinclair, M., G. Buhrmann, et B. Gummow. 2006. « An epidemiological investigation of the African horsesickness outbreak in the Western Cape Province of South Africa in 2004 and its relevance to the current equine export protocol ». *Journal of the South African Veterinary Association* 77 (4): 191-96. doi:10.4102/jsava.v77i4.376.
- Smithburn, K. C., T. P. Hughes, A. W. Burke, et J. H. Paul. 1940. « A Neurotropic Virus Isolated from the Blood of a Native of Uganda ». *The American Journal of Tropical Medicine and Hygiene* 1-20 (4): 471-92.
- Smith, Kristine M., Simon J. Anthony, William M. Switzer, Jonathan H. Epstein, Tracie Seimon, Hongwei Jia, Maria D. Sanchez, et al. 2012. « Zoonotic Viruses Associated with Illegally Imported Wildlife Products ». *PLoS ONE* 7 (1): e29505. doi:10.1371/journal.pone.0029505.
- Sonesson, Christian, et Marianne Frisé. 2005. « Multivariate Surveillance ». In *Spatial and Syndromic Surveillance for Public Health*, édité par Andrew B. Lawson et Ken Kleinman, 153-66. John Wiley & Sons, Ltd. <http://onlinelibrary.wiley.com/doi/10.1002/0470092505.ch9/summary>.
- Sotelo, Elena, Ana Gutierrez-Guzmán, Javier del Amo, Francisco Llorente, Mehdi El-Harrak, Elisa Pérez-Ramírez, Juan Blanco, Ursula Höfle, et Miguel Jiménez-Clavero. 2011. « Pathogenicity of two recent Western Mediterranean West Nile virus isolates in a wild bird species indigenous to Southern Europe: the red-legged partridge ». *Veterinary Research* 42 (1): 1-8. doi:10.1186/1297-9716-42-11.
- Stärk, Katharina DC, Gertraud Regula, Jorge Hernandez, Lea Knopf, Klemens Fuchs, Roger S Morris, et Peter Davies. 2006. « Concepts for risk-based surveillance in the field of veterinary medicine and veterinary public health: Review of current approaches ». *BMC Health Services Research* 6 (février): 20. doi:10.1186/1472-6963-6-20.
- Steele, K. E., M. J. Linn, R. J. Schoepp, N. Komar, T. W. Geisbert, R. M. Manduca, P. P. Calle, et al. 2000. « Pathology of Fatal West Nile Virus Infections in Native and Exotic Birds during the 1999

- Outbreak in New York City, New York ». *Veterinary Pathology Online* 37 (3): 208-24.
doi:10.1354/vp.37-3-208.
- Stevens, K. B., S. Costard, R. Métras, et D. U. Pfeiffer. 2009. « Mapping the Likelihood of Introduction and Spread of Highly Pathogenic Avian Influenza Virus H5N1 in Africa, Ghana, Ethiopia, Kenya and Nigeria using Multicriteria Decision Modelling ». *DFID funded project for Controlling Avian Flu and Protecting People's Livelihoods in Africa/Indonesia*.
http://www.ifpri.org/sites/default/files/publications/hpaiwp_africa.pdf.
- Stoto, Michael A., Ronald D. Fricker Jr, Arvind Jain, Alexis Diamond, John O. Davies-Cole, Chevelle Glymph, Gebreyesus Kidane, Garrett Lum, LaVerne Jones, et Kerda Dehan. 2006. « Evaluating statistical methods for syndromic surveillance ». In *Statistical Methods in Counterterrorism: Game Theory, Modeling, Syndromic Surveillance, and Biometric Authentication*, 141-72. Springer. doi:10.1007/0-387-35209-0_9. http://link.springer.com/content/pdf/10.1007/0-387-35209-0_9.pdf.
- Stroup, Donna F., G. David Williamson, Joy L. Herndon, et John M. Karon. 1989. « Detection of Aberrations in the Occurrence of Notifiable Diseases Surveillance Data ». *Statistics in Medicine* 8 (3): 323-29.
doi:10.1002/sim.4780080312.
- Taroni, Franco, Colin Aitken, Paolo Garbolino, et Alex Biedermann. 2006. *Bayesian Networks and Probabilistic Inference in Forensic Science*. Statistics in Practice. Chichester, England: Wiley-Blackwell.
- Theiler, Arnold F. 1910. « Notes on a fever in horses simulating horse-sickness. », *Transvaal Agricultural Journal*, 8 (32): 581-86.
- Tran, Annelise, Bertrand Sudre, Shlomit Paz, Massimiliano Rossi, Annie Desbrosse, Véronique Chevalier, et Jan C. Semenza. 2014. « Environmental predictors of West Nile fever risk in Europe ». *International journal of health geographics* 13 (1): 26. doi:10.1186/1476-072X-13-26.
- Triple S Project. 2011. « Guideline for designing and implementing a syndromic surveillance system ». Guidelines. http://www.syndromicsurveillance.eu/Triple-S_guidelines.pdf.
- Trock, Susan C., Barry J. Meade, Amy L. Glaser, Eileen N. Ostlund, Robert S. Lanciotti, Bruce C. Cropp, Varuni Kulasekera, Laura D. Kramer, et Nicholas Komar. 2001. « West Nile virus outbreak among horses in New York State, 1999 and 2000. » *Emerging infectious diseases* 7 (4): 745.
doi:10.3201/eid0704.010427.
- Tsui, F. C., M. M. Wagner, V. Dato, et C. C. Chang. 2001. « Value of ICD-9 coded chief complaints for detection of epidemics. » In *Proceedings of AMIA Annual Symposium*, 711-15.
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2243339/>.
- Turner, Joanne, Roger G Bowers, et Matthew Baylis. 2013. « Two-Host, Two-Vector Basic Reproduction Ratio ($R(0)$) for Bluetongue ». *PLoS One* 8 (1): e53128. doi:10.1371/journal.pone.0053128.
- USDA. 2012. « Guidelines for Plant Pest Risk Assessment of Imported Fruit & Vegetable Commodities ». Guidelines. USDA, Animal and Plant Health Inspection Service, Plant Protection and Quarantine.
http://www.aphis.usda.gov/import_export/plants/plant_imports/process/downloads/PRAGuidelines-ImportedFruitVegCommodities.pdf.
- Valenciano, M. 2001. « Définition des priorités dans le domaine des zoonoses non-alimentaires ». Paris: Institut de Veille Sanitaire.
http://www.invs.sante.fr/publications/2010/zoonose_non_alimentaire/rapport_zoonoses_non_alimentaires.pdf.
- Valiakos, George, Konstantinos Papaspyropoulos, Alexios Giannakopoulos, Periklis Birtsas, Sotirios Tsiodras, Michael R. Hutchings, Vassiliki Spyrou, et al. 2014. « Use of Wild Bird Surveillance, Human Case Data and GIS Spatial Analysis for Predicting Spatial Distributions of West Nile Virus in Greece ». *PLoS ONE* 9 (5): e96935. doi:10.1371/journal.pone.0096935.

- Van der Meulen, K. M., M. B. Pensaert, et H. J. Nauwynck. 2005. « West Nile virus in the vertebrate world ». *Archives of Virology* 150 (4): 637-57. doi:10.1007/s00705-004-0463-z.
- Veksler, Anna, Millicent Eidson, et Igor Zurbenko. 2009. « Assessment of methods for prediction of human West Nile virus (WNV) disease from WNV-infected dead birds ». *Emerging Themes in Epidemiology* 6 (1): 1-10. doi:10.1186/1742-7622-6-4.
- Venter, G. J., S. D. Graham, et C. Hamblin. 2000. « African Horse Sickness Epidemiology: Vector Competence of South African Culicoides Species for Virus Serotypes 3, 5 and 8 ». *Medical and Veterinary Entomology* 14 (3): 245-50. doi:10.1046/j.1365-2915.2000.00245.x.
- Venter, G. J., D. M. Groenewald, J. T. Paweska, E. H. Venter, et P. G. Howell. 1999. « Vector Competence of Selected South African Culicoides Species for the Bryanston Serotype of Equine Encephalosis Virus ». *Medical and Veterinary Entomology* 13 (4): 393-400. doi:10.1046/j.1365-2915.1999.00188.x.
- Venter, G. J., D. Groenewald, E. Venter, K. G. Hermanides, et P. G. Howell. 2002. « A Comparison of the Vector Competence of the Biting Midges, Culicoides (Avaritia) Bolitinos and C. (A.) Imicola, for the Bryanston Serotype of Equine Encephalosis Virus ». *Medical and Veterinary Entomology* 16 (4): 372-77. doi:10.1046/j.1365-2915.2002.00385.x.
- Venter, G. J., I. M. Wright, et J. T. Paweska. 2010. « A comparison of the susceptibility of the biting midge Culicoides imicola to infection with recent and historical isolates of African horse sickness virus ». *Medical and veterinary entomology* 24 (3): 324-28. doi:10.1111/j.1365-2915.2010.00895.x.
- Vial, Flavie, et John Berezowski. 2014. « A practical approach to designing syndromic surveillance systems for livestock and poultry ». *Preventive Veterinary Medicine*. Consulté le décembre 12. doi:10.1016/j.prevetmed.2014.11.015.
- Viennet, E., C. Garros, L. Gardès, I. Rakotoarivony, X. Allène, R. Lancelot, D. Crochet, C. Moulia, T. Baldet, et T. Balenghien. 2013. « Host Preferences of Palaearctic Culicoides Biting Midges: Implications for Transmission of Orbiviruses ». *Medical and Veterinary Entomology* 27 (3): 255-66. doi:10.1111/j.1365-2915.2012.01042.x.
- Viljoen, G. J., et H. Huismans. 1989. « The Characterization of Equine Encephalosis Virus and the Development of Genomic Probes ». *Journal of General Virology* 70 (8): 2007-15. doi:10.1099/0022-1317-70-8-2007.
- Vose, David. 2008. *Risk Analysis: A Quantitative Guide*. 3rd Edition. Chichester, England ; Hoboken, NJ: John Wiley & Sons.
- Ward, Michael P., James A. Schuermann, Linda D. Highfield, et Kristy O. Murray. 2006. « Characteristics of an outbreak of West Nile virus encephalomyelitis in a previously uninfected population of horses ». *Veterinary Microbiology* 118 (3-4): 255-59. doi:10.1016/j.vetmic.2006.07.016.
- Webster, W. R. 2011. « Overview of the 2007 Australian Outbreak of Equine Influenza ». *Australian Veterinary Journal* 89 (juillet): 3-4. doi:10.1111/j.1751-0813.2011.00721.x.
- Wescott, David G., Zvia Mildenberg, Michel Bellaiche, Sarah L. McGowan, Sylvia S. Grierson, Bhudipa Choudhury, et Falko Steinbach. 2013. « Evidence for the Circulation of Equine Encephalosis Virus in Israel since 2001 ». *PLoS ONE* 8 (8): e70532. doi:10.1371/journal.pone.0070532.
- Wilson, Anthony, Philip Scott Mellor, Camille Szmaragd, et Peter Paul Clement Mertens. 2009. « Adaptive strategies of African horse sickness virus to facilitate vector transmission ». *Veterinary Research* 40 (2). doi:10.1051/vetres:2008054.
- Wittmann, E. J., P. S. Mellor, et M. Baylis. 2002. « Effect of Temperature on the Transmission of Orbiviruses by the Biting Midge, Culicoides sonorensis ». *Medical and Veterinary Entomology* 16 (2): 147-56. doi:10.1046/j.1365-2915.2002.00357.x.

- Yu, Peter, Tsegaye Habtemariam, Saul Wilson, David Oryang, David Nganwa, Mike Obasa, et Vinaida Robnett. 1997. « A risk-assessment model for foot and mouth disease (FMD) virus introduction through deboned beef importation ». *Preventive Veterinary Medicine* 30 (1): 49-59. doi:10.1016/S0167-5877(96)01085-9.
- Zeller, H. 2010. « West Nile Virus: The Need to Strengthen Preparedness in Europe », *Eurosurveillance*, 15 (34). <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=19647>.
- Ziegler, Ute, Joke Angenvoort, Dominik Fischer, Christine Fast, Martin Eiden, Ariel V. Rodriguez, Sandra Revilla-Fernández, et al. 2013. « Pathogenesis of West Nile Virus Lineage 1 and 2 in Experimentally Infected Large Falcons ». *Veterinary Microbiology* 161 (3-4): 263-73. doi:10.1016/j.vetmic.2012.07.041.
- Zientara, Stéphan, et José Manuel Sánchez-Vizcaíno. 2013. « Control of bluetongue in Europe ». *Veterinary Microbiology* 166 (1-2): 33-37. doi:10.1016/j.vetmic.2013.01.010.
- Zimmerli, U, C Herholz, H Schwermer, M Hofmann, et C Griot. 2010. « [African horse sickness and equine encephalosis: must Switzerland get prepared] ». *Schweizer Archiv für Tierheilkunde* 152 (4): 165-75. doi:10.1024/0036-7281/a000039.
- Zou, Jian, Alan F. Karr, David Banks, Matthew Heaton, Gauri Datta, James Lynch, et Francisco Vera. 2010. « Bayesian methodology for spatio-temporal syndromic surveillance ». Technical report 174. NC, USA: National Institute of Statistical Sciences. <http://www.niss.org/sites/default/files/tr174.pdf>.

ABSTRACT

Emerging vector-borne diseases are a growing concern, especially for horse populations, which are at particular risk for disease spread. In general, horses travel widely and frequently and, despite the health and economic impacts of equine diseases, effective health regulations and biosecurity systems to ensure safe equine movements are not always in place. The present work proposes to improve the surveillance of vector-borne diseases in horses through the use of different approaches that assess the probability of occurrence of a newly introduced epidemic. First, we developed a spatiotemporal quantitative model which combined various probabilities in order to estimate the risk of introduction of African horse sickness and equine encephalosis. Such combinations of risk provided more a detailed picture of the true risk posed by these pathogens. Second, we assessed syndromic surveillance systems using two approaches: a classical approach with the alarm threshold based on the standard error of prediction, and a Bayesian approach based on a likelihood ratio. We focused particularly on the early detection of West Nile virus using reports of nervous symptoms in horses. Both approaches provided interesting results but Bayes' rule was especially useful as it provided a quantitative output and was able to combine different epidemiological information. Finally, a Bayesian approach was also used to quantitatively combine various sources of risk estimation in a multivariate syndromic surveillance system, as well as a combination of quantitative risk assessment with syndromic surveillance (applied to West Nile virus and equine encephalosis, respectively). Combining evidence provided promising results. This work, based on risk estimations, strengthens the surveillance of VBDs in horses and can support public health decision making. It also, however, highlights the need to improve data collection and data sharing, to implement full performance assessments of complex surveillance systems, and to use effective communication and training to promote the adoption of these approaches.

Key words: syndromic surveillance, West Nile, quantitative risk analysis, African horse sickness, equine encephalosis, vector-borne diseases, risk-based surveillance

RÉSUMÉ

Les maladies émergentes à transmission vectorielle sont une préoccupation croissante et particulièrement lorsqu'elles affectent les chevaux, une population spécifiquement à risque vis-à-vis de la propagation de maladies. En effet, les chevaux voyagent fréquemment et, malgré l'impact sanitaire et économique des maladies équine, les réglementations sanitaires et les principes de biosécurité et de traçabilité censés assurer la sécurité des mouvements d'équidés ne sont pas toujours en place. Notre travail propose d'améliorer la surveillance des maladies à transmission vectorielle chez les chevaux en utilisant différentes méthodes pour estimer la probabilité d'émergence d'une maladie. Tout d'abord, nous avons développé un modèle quantitatif et spatio-temporel combinant différentes probabilités pour estimer les risques d'introduction de la peste équine et de l'encéphalose équine. Ces combinaisons permettent d'obtenir une image plus détaillée du risque posé par ces agents pathogènes. Nous avons ensuite évalué des systèmes de surveillance syndromique par deux approches méthodologiques: l'approche classique avec un seuil d'alarme basé sur un multiple de l'erreur standard de prédiction, et l'approche bayésienne basée sur le rapport de vraisemblance. Nous avons travaillé ici principalement sur la détection précoce du virus West Nile en utilisant les symptômes nerveux des chevaux. Les deux approches ont fourni des résultats prometteurs, mais l'approche bayésienne était particulièrement intéressante pour obtenir un résultat quantitatif et pour combiner différentes informations épidémiologiques. Pour finir, l'approche bayésienne a été utilisée pour combiner quantitativement différentes sources d'estimation du risque : surveillance syndromique multivariée, et combinaison de la surveillance syndromique avec les résultats d'analyses de risques. Ces combinaisons ont données des résultats prometteurs. Ce travail, basé sur des estimations de risque, contribue à améliorer la surveillance des maladies à transmission vectorielle chez les chevaux et facilite la prise de décision. Les principales perspectives de ce travail sont d'améliorer la collecte et le partage de données, de mettre en œuvre une évaluation complète des performances des systèmes de surveillance multivariés, et de favoriser l'adoption de ce genre d'approche par les décideurs en utilisant une interface conviviale et en mettant en place un transfert de connaissance.

Mots clefs : surveillance syndromique, West Nile, analyse de risques quantitative, peste équine, encéphalose équine, maladies à transmission vectorielle, surveillance basée sur le risque