SOCIETY FOR VETERINARY EPIDEMIOLOGY AND PREVENTIVE MEDICINE

Proceedings of a meeting held in

Madrid, Spain

20th – 22nd March 2013

Edited by K.L.P. Verheyen and C. Fourichon
and the SVEPM Executive Committee

ISBN 978-0-948073-20-5

ACKNOWLEDGEMENTS

The following bodies provided financial support for the conference and the publication of these proceedings:

CIRAD

COLVEMA
(Colegio Oficial de Veterinarios de Madrid)

Pfizer Animal Health

Nordic Society for Veterinary Epidemiology

Ministerio de Agricultura, Alimentación y Medio Ambiente
(Ministry of Agriculture, Food and Environment, Spain)

Avia-GIS

The Royal Veterinary College

EpiX Analytics

CONTENTS

FOOT AND MOUTH DISEASE

DISEASE TRANSMISSION/CONTROL

VETERINARY PUBLIC HEALTH

SYNDROMIC SURVEILLANCE

OPEN SESSION

OBSERVATIONAL STUDIES

FARMERS' PERCEPTIONS AND DISEASE CONTROL

METHODS

***

Society for Veterinary Epidemiology and Preventive Medicine

Past Venues & Conference Organisers
Past Presidents
Executive Committee Members & Life Members
Gareth Davies Lectures & Opening Conference Plenary Lectures
Membership Application Form & Treasurer’s Address
Constitution and Rules

FOOT AND MOUTH DISEASE

DECISIONS ON FOOT-AND-MOUTH DISEASE CONTROL INFORMED BY MODEL PREDICTIONS

T. HALASA*, P. WILLEBERG, L.E. CHRISTIANSEN, A. BOKLUND, M. ALKHAMIS, A. PEREZ AND C. ENØE

SUMMARY

The predictive capability of the first fortnight incidence (FFI), which is the number of detected herds within the first 14 days following detection of the disease, of the course of a foot-and-mouth disease (FMD) epidemic and its outcomes were investigated. Epidemic outcomes included the number of affected herds, epidemic duration, geographical size, and costs. The first fourteen days spatial spread (FFS) was also included to support the prediction. The epidemic data were obtained from a Danish version (DTU-DADS) of the Davis Animal Disease Spread simulation model.

The FFI and FFS showed good correlations with the epidemic outcomes. The predictive capability of the FFI was high. This indicates that the FFI may take a part in the decision of whether or not to boost FMD control, which might prevent occurrence of a large epidemic in the face of an FMD incursion. The prediction power was improved by supplementing the models with information on FFS and characteristics of the index-herd. Results presented here will contribute to improve preparedness of Denmark to early control of a hypothetical FMD epidemic.

INTRODUCTION

Foot-and-mouth disease (FMD) is a highly contagious viral disease affecting ruminants and pigs that may cause a large economic damage to affected countries (Pendell et al., 2007). From an ethical point of view, protective emergency vaccination has the advantage of decreasing the number of slaughtered animals during an epidemic (Hutber et al., 2006). Furthermore, the World Organization for Animal Health (OIE) and the European Union (EU) have revised their regulations to enable use of emergency vaccination by reducing the time delay to regain FMD-free status (European Commission, 2003; OIE, 2011), and thus resume export sooner than in the past, which may reduce potentially large economic damage. Nonetheless, protective vaccination is still a dubious option for EU countries with large export of livestock and livestock products. A large associated economic damage may be expected, due to unpredictable delays of non-EU countries to resume imports from the affected EU member states. Therefore, the decision on whether or not to apply emergency

---

* Tariq Halasa, The National Veterinary Institute, Technical University of Denmark, Bülowsvej 27, 1870 Frederiksberg C, Denmark. Email: tahbh@vet.dtu.dk

vaccination is probably one of the most difficult decisions facing the veterinary authorities in cases of an FMD outbreak (Barnett et al., 2002).

Based on regional analysis of outbreak data from the FMD outbreak that affected the UK in 2001, Hutber et al. (2006) proposed to use the first fortnight incidence (FFI, i.e., the number of FMD detected herds during the first 14 days of an epidemic) to estimate the epidemic outcome. The number of affected herds per region by the end of the epidemic was the only epidemic outcome included in that early study (Hutber et al., 2006). However, other epidemic outcomes that have substantial economic consequences, such as epidemic duration and the geographical size of the affected area ("epidemic geo-size"), were not examined. In the current study we used output data from a pre-existing FMD simulation model (Bates et al., 2003) adapted to simulate the spread of FMD in Denmark to quantify the ability of the FFI and/or the first fourteen days spatial spread (FFS) of FMD (i.e., the circumference of the detected area at day 14 of the epidemic) to predict the course of the epidemic. The output parameters included the number of affected herds, the epidemic duration, the epidemic geo-size, and the economic consequences of the epidemic. A good predictive power of such early estimates would give decision makers a tool to assist them in their difficult but important considerations of whether or not to step up the control strategy to include pre-emptive depopulation or emergency vaccination of susceptible herds.

MATERIALS AND METHODS

The simulation study

The Davis Animal Disease Simulation model (DADS, version 0.05) (Bates et al., 2003) was further adapted to the Danish setting (DTU-DADS, version 0.100) to model the spread of FMD in Denmark, using actual herd data and input values that are described in detail elsewhere (Boklund et al., unpublished). Herd data included the Danish Herd Identification System, referred to as CHR number, herd type (cattle, swine, and sheep), UTM geo-coordinates, number of animals per age group, and the rate of animal movements per day. When a farm had more than one species, each species was considered a separate herd, but located on the same location. Animal movements were categorised as herd-to-herd and herd-to-abattoir movements. Information about markets, including the UTM coordinates, was also included.

Each epidemic started at one randomly chosen herd, which was referred to as the index-herd. To assess for the effect of variation in the type and location of the index-herd, epidemics were initiated in 1,000 cattle herds located in a high density cattle area and 1,000 in a low density cattle area, 1,000 swine herds located in a high density swine area and 1,000 in a low density swine area, and 1,000 sheep herds. The model was run once (i.e., one iteration) per index-herd.

Spread of infection amongst herds was simulated through 7 spread mechanisms: (1) direct animal movement between herds, (2) abattoir truck, (3) milk tanker, (4) veterinarians, artificial inseminators, and/or milk controllers (medium risk contact), (5) visitors, feedstuff and/or rendering trucks (low risk contacts), (6) markets, and (7) local spread.

Based on actual movement data, the number of movements per day was calculated for each herd, representing a daily rate of outgoing movement per herd. This value was assumed to be the mean (λ_h) in a Poisson distribution. Similarly, the rate of abattoir deliveries per day was calculated based on actual herd data and used as the mean (λ_a) in a Poisson distribution. A Poisson distribution with a mean value (λ_m) of 0.6 was used to represent how often milk is picked up in dairy herds. Likewise, medium and low risk contacts were simulated (Boklund et al., unpublished).

Because markets in Denmark are restricted to cattle only, an infection spreading from a market can initially affect only cattle herds. A movement of infected animals to a market was estimated to result in an average of 3.5 extra contacts (Boklund et al., unpublished). The spread via markets would be due to direct movements of infected animals to susceptible herds, or via people and vehicles that had been in contact with the infected animals and then in contact with susceptible herds. Local spread was defined as infection of susceptible herds within a 3 km radius around the infected herd due to unknown reasons. Farms may consist of more than one herd. Herds located on the same farm had a daily chance of infection of 95%, when one herd was infected. Per definition, the disease always started in the index-herd and developed until the disease was detected and hence the herd was depopulated. The period from a herd starting showing clinical signs until detection was dependent on the herd type, e.g. cattle herds were detected faster than sheep herds. Moreover, herds within the protection and surveillance zones would have higher probability of detection, because of surveillance.

For simplicity, detection of the first infected herd was assumed to occur at day 21 following the introduction of the FMD virus. This was based on experience from the UK (Gibbens et al., 2001; Gibbens & Wilesmith, 2002) and the Dutch 2001 FMD outbreaks (Pluimers et al., 2002). In the simulation, the infection spreads freely between herds during the first 21 days. After detection of the first infected herd, a set of control strategies were applied. These included: (1) depopulation, cleaning and disinfection of detected herds; (2) a 3 days national stand still on animal movements in the country; (3) a 10 km radius zone (surveillance zone) around the detected herds for 30 days, in which movements between herds and out of the zone were restricted and herds were surveyed one time before lifting the zone; (4) a 3 km radius zone (protection zone) around the detected herd for 21 days, in which movements between herds and out of the zone were restricted, and herds were inspected during the first week; (5) backward and forward tracing of contacts from and to detected herds. Herds that received animals from a detected herd were depopulated and disinfected, whereas in case of other kinds of contact, the animals were inspected for clinical signs of FMD. In case of sheep herds, the animals were also sampled for serological analysis. The daily animal depopulation capacity was 2,400 ruminants and 4,800 pigs. Detected herds had higher priority for depopulation than contact herds. In case of several herds on the same farm, all herds on the farm were depopulated when one herd was detected.

The total costs of the epidemic were computed for each model iteration as the sum of the direct and indirect costs associated with the epidemic. The direct costs were due to surveillance, depopulation, cleaning and disinfection, empty stables, compensation, and national standstill costs. The indirect costs included losses incurred from restrictions on trade with EU and non-EU countries. The FFS and the epidemic geo-size were calculated by plotting the locations of the detected herds per iteration and constructing a minimum convex hull around the herds and then measuring the circumference of the resulting convex hull polygon (Eddy, 1977).

Statistical analysis

The FFI, FFS and epidemic outcomes including cumulative number of affected herds, epidemic duration, geo-size, and costs (€) were calculated per iteration and gathered in a dataset consisting of 5,000 iterations. Data processing and analysis was carried out in R 2.14.0 (R development core team, 2011). The correlation between the FFI and FFS and the epidemic outcomes was obtained using Spearman's correlation coefficient.

Linear regression models were fitted to investigate the prediction capability of the FFI and FFS and correcting for index-herd type according to Eq. (1).

$$(Y)^z = \text{Intercept} + \text{FFI} + \text{FFS} + \text{FFI} \times \text{FFS} + \text{Index-herd type} \quad (1)$$

Y was the additional number of affected herds after day 14, epidemic duration after day 14, epidemic geo-size $\times 10^{-3}$ or epidemic costs $\times 10^{-9}$. The Box-Cox power transformation method (Box & Cox, 1964) was used to transform Y to the power value z, in order to stabilise the variance and to make the residuals best fitting the normal distribution. Index-herd type was a categorical variable consisting of 5 groups representing the 5 different index-herd types presented above. Interaction terms between index-herd type and the other independent variables were also examined. The FFI was transformed using the square root to make the residuals best fit the normal distribution, while the FFS was multiplied by 10^{-3}. Furthermore, when epidemic geo-size was the dependent variable (Y), FFS was further transformed to the power -2 for numerical stabilisation. The models were fitted using the backward selection method and model assumptions including residuals' normality and homogeneity were confirmed.

RESULTS

The median and (5th and 95th percentiles) of the FFI and FFS were 20 (4-67) herds and 242 (22-466) km, respectively. The median and (5th and 95th percentiles) of epidemic outcomes per index-herd type are presented in Table 1. The table shows that epidemics started in cattle herds would last longer, be larger, and cost more than epidemics started in swine or sheep herds. Out of the 5,000 simulated epidemics, 4,458 lasted > 14 days and were used in the subsequent analyses.

Table 1. Median and (5th and 95th percentiles) of epidemic outcomes including epidemic duration (in days), number of infected herds, the geo-size (in km) and the total costs (in €×10^6) of the epidemic that started in cattle herds located in high or low cattle density areas, in swine herds located in high or low swine density areas, and in sheep herds

Index-herd type	Epidemic duration	Infected herds	Geo-size	Costs
High cattle	56 (16-142)	67 (13-245)	412 (129-678)	562 (402-946)
Low cattle	71 (19-179)	94 (15-371)	450 (78-748)	604 (416-1061)
High swine	43 (8-130)	36 (5-195)	322 (30-669)	502 (376-869)
Low swine	43 (7-147)	36 (4-224)	317 (28-652)	493 (373-904)
Sheep	38 (6-139)	29 (3-198)	290 (5-636)	479 (364-875)

The FFI and FFS showed a fairly good correlation with epidemic outcomes (Table 2). The highest correlation obtained for the FFI and the epidemic outcomes was with epidemic geo-size. The FFS had generally higher correlation with epidemic outcomes than the FFI (Table 2).

Table 2. Spearman correlation coefficient between the first fortnight incidence (FFI) and the first fourteen days spatial spread (FFS) of the epidemic and epidemic outcomes including the additional number of affected herds after day 14, epidemic duration after day 14, geo-size and costs

Epidemic outcomes	FFI	FFS
Additional number of affected herds after day 14	0.66	0.82
Epidemic duration after day 14	0.50	0.76
Geo-size	0.69	0.70
Costs	0.59	0.82

Epidemic outcome parameters were power transformed so that the residuals would best fit the normal distribution and to stabilise the variance. For instance, the additional number of affected herds was transformed to the power value 0.16. The outcomes of Eq. (1) fitted to the additional number of affected herds after day 14 are presented in Table 3.

Table 3. Parameter estimates together with the standard error (s.e.) in brackets for the model fitted to explain the additional number of affected herds after day 14, using the first fortnight incidence (FFI) together with the first fourteen days spatial spread (FFS) of the epidemic and the index-herd type

Epidemic outcomes	Coefficient value with (s.e.)
Intercept	0.83 (0.03)
$FFI^{0.5}$	0.16 (0.006)
FFS	0.79 (0.09)
$FFI^{0.5}$*FFS	-0.096 (0.017)
Index-Herd type	
Low cattle	0.09 (0.013)
High swine	0.06 (0.014)
Low swine	0.07 (0.014)
Sheep	0.08 (0.015)

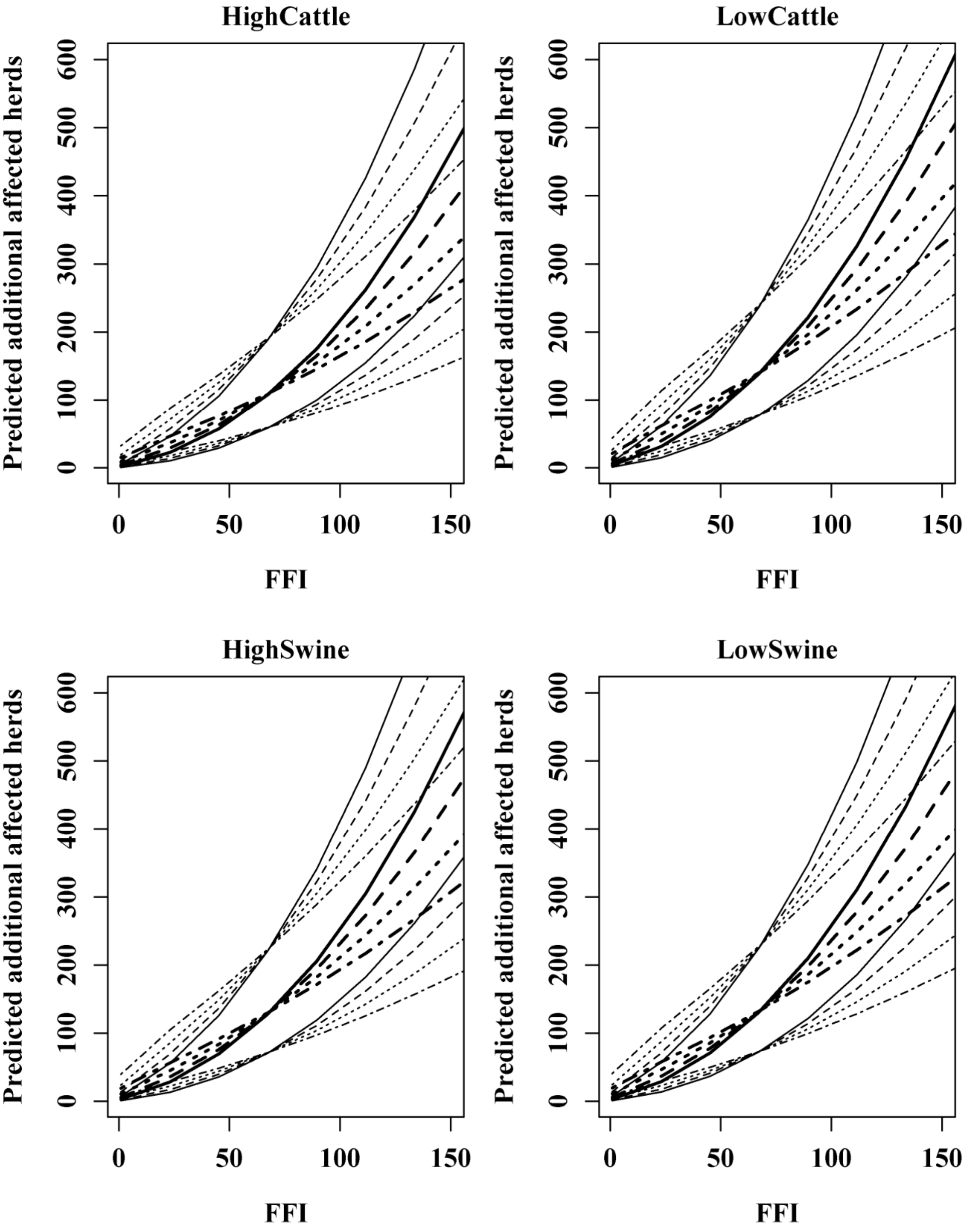


Fig. 1 Predicted additional number of affected herds given the first fortnight incidence (FFI) for four categories of first fourteen days spatial spread of the epidemic (FFS) and when epidemics started in cattle herds located in a high density cattle area (HighCattle), in a low density cattle area (LowCattle), in swine herds located in a high density swine area (HighSwine) and in a low swine density area (LowSwine). The FFS categories were ≤ 200 km, 200 to 400 km, 400 to 600 km, and > 600 km and were represented by the continuous, dashed, dotted and dot-dash lines respectively. The thick lines represent the median predictions and the corresponding thin lines represent the interquartile range

To visually illustrate the outcomes of Table 3, the predicted number of additional affected herds after day 14 for four index-herd types and FFS categories that were made based on histogram categories, following different FFI values are presented in Fig.1. The figure shows that the predicted additional number of affected herds increases exponentially with FFI values. Furthermore, it shows that the effect of FFI changes corresponding to changes to the FFS, reflecting the interaction between FFI and FFS. This interaction takes effect when the FFI is between 60 and 70 detected herds, depending on the index-herd type. Figure 2 shows the median predicted economic consequence of for different FFI values and FFS categories, and when the epidemics started in cattle herds located in high density cattle areas.

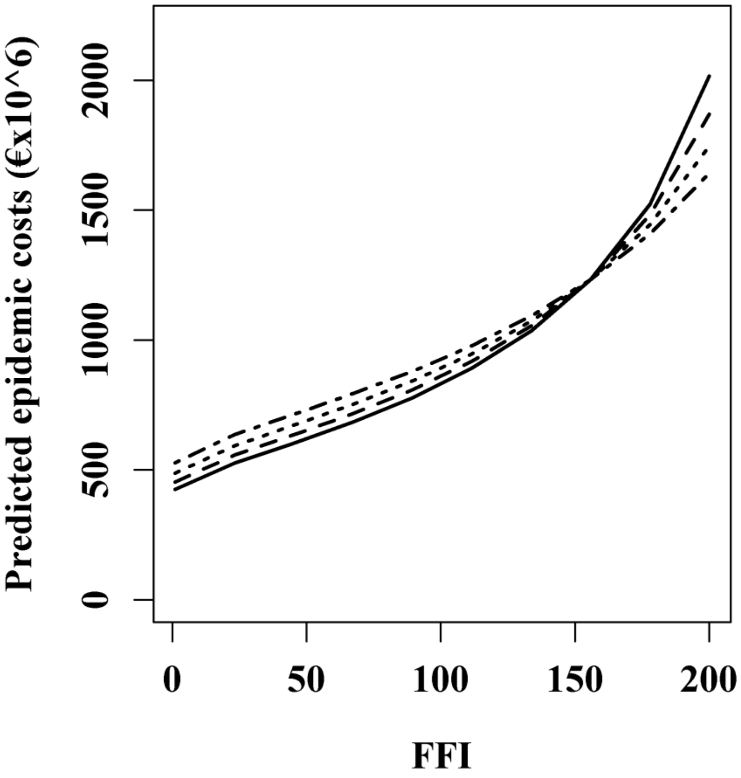


Fig. 2 Predicted epidemic costs in €×10^6 given the first fortnight incidence (FFI) for four categories of first fourteen days spatial spread of the epidemic (FFS) and when epidemics started in cattle herds located in high density cattle area. The FFS categories were ≤ 200 km, 200 to 400 km, 400 to 600 km, and > 600 km and were represented by the continuous, dashed, dotted and dot-dash lines respectively

DISCUSSION

The number of detected herds within the first 14 days is rather important in assessing the prognosis of the whole outbreak and its economic consequences. The larger the number of affected herds at day 14, the higher the additional number of affected herds and the epidemic costs. This information provides the veterinary authorities the chance to reconsider the applied FMD control strategies 14 days following the detection of the epidemic by, for instance, introducing emergency vaccination that has the potential to prevent catastrophic situations (Hagerman et al., 2010).

FFI was correlated with and a good predictor of the additional number of affected herds after day 14, which is consistent with results reported by Hutber et al. (2006). Nevertheless, the correlation in this study was slightly lower than that observed by Hutber et al. (2006). At day 14 following the first detection of the outbreak, the veterinary authorities would most likely know the exact number of detected herds, their locations, and the index-herd. This information can be inserted into the model presented in Table 3 to predict the additional number of affected herds following day 14, in an easy and direct way, which would be rather useful for the veterinary authorities. Likewise, the economic consequence can be predicted as presented in Fig. 2.

The interaction between FFI and FFS (Fig.1) suggests that the effect of FFI will change depending on changes to FFS. The procedure of examining interaction between two continuous variables by Dohoo et al. (2003) was followed, and FFS was categorised based on histogram categories. Thereafter, the predicted values of additional number of affected herds after day 14 were plotted against the FFI per category (Fig.1). The figure shows that when the FFI < approximately 65 herds that are spread in a small area (the continuous line), the additional number of affected herds would be small, compared to had the detected herds (within the first 14 days) been spread over a larger area. Nonetheless, when the FFI > approximately 65 herds, the effect will be opposite. This interaction may be explained by the effect of control strategies, and because the FFS and FFI are correlated.

Contradictory to Hutber et al. (2006), who conducted regional analysis on outbreak data, the data in the current study were obtained from a simulation model, in which each iteration represents a different epidemic. This gave us the chance to test the effects on wide range of different epidemics. This is not possible using outbreak data, as outbreaks are usually specific. Nonetheless, the behaviour of the model is usually highly dependent on input parameters, which requires model validation. External validation of models is usually carried out by comparing model outputs to actual data, and in this case outbreak data. Denmark has not, however, experienced FMD outbreaks in the past 30 years, which makes it impossible to directly validate the model data, and leaves simulation modelling as the only option to study FMD spread in Denmark. Our reliance on the value of the simulation data comes from the fact that the model simulated the spread of FMD between herds using actual and current herd movement and contact data. Thus this study complements the work of Hutber et al. (2006) and provides a useful tool that can be applied to other countries. It is important to mention, however, that the predictors and their effects identified in this study are country-specific and would differ in magnitude between countries, because of differences in herd structure, farming systems, management, regulations and national priorities and preferences. Results presented here will improve the ability of Denmark to control an outbreak early and, ultimately, mitigate the impact of a hypothetical FMD incursion into the country.

ACKNOWLEDGEMENTS

This study was financially supported by the Directorate for Food, Agriculture and Fisheries, Denmark (grant nr. 3304-FVFP-07-782-01).

REFERENCES

Barnett, P., Garland, A.J.M., Kitching, R.P. and Schermbruker, C.G. (2002). Aspects of emergency vaccination against FMD. Comp. Immunol. Microb. 25, 345-364

Bates, T., Thurmond, M.C. and Carpenter, T.E. (2003). Description of and epidemic simulation model for use in evaluating strategies to control an outbreak of foot-and-mouth disease. Am. J. Vet. Res. 64, 195-204

Box, G.E.P. and Cox, D.R. (1964). An analysis of transformations. J. Roy. Stat. Soc. B 26, 211-252

Dohoo, I.R., Martin, S.W. and Stryhn, H. (2003). Veterinary Epidemiologic Research. Atlantic Veterinary College Inc., Charlottetown, Prince Edward Island, Canada

Eddy, W.F. (1977). A new convex hull algorithm for planar **sets.** ACM Trans. Math. Soft. 3, 398-403

European Commission (2003). Council directive 03/85/EC on community measures for the control of foot-and-mouth disease repealing (2003) Directive 85/511/EEC and amending directive 92/46/EEC. Official Journal of European Union L306, 46, 22 November 2003

Hagerman, A.D., Mccarl, B.A. and Lin, H.I. (2010). The role of risk attitude in preference rankings of vaccine use for FMD eradication in the U.S., Poster prepared for presentation at the Agricultural & Applied Economics Association 2010, AAEA, CAES, & WAEA Joint Annual Meeting, Denver, Colorado, July 25-27

Hutber, A.M., Kitching, R.P. and Phillipcinec, E. (2006). Predictions for the timing and use of culling or vaccination during a foot-and-mouth disease epidemic. Res. Vet. Sci. 81, 31-36

Gibbens, J.C., Share, C.E., Wilesmith, J.W., Mansley, L.M., Michalopoulou, E., Ryan, J.B.M. and Hudson, M. (2001). Descriptive epidemiology of the 2001 foot-and-mouth disease epidemic in Great Britain: the first five months. Vet. Rec. 149, 729-743

Gibbens, J.C. and Wilesmith, J.W. (2002). Temporal and geographical distribution of the cases of foot-and-mouth disease during the early weeks of the 2001 epidemic in Great Britain. Vet. Rec. 151, 407-412

Pendell, D.L., Leatherman, J., Schroeder, T.C. and Alward, G.S. (2007). The economic impact of foot-and-mouth disease outbreak: a regional analysis. J. Agr. Appl. Econ. 39, 19-33

Pluimers, F.H., Akkerman, A.M., Van Der Wal, P., Dekker, A. and Bianchi, A. (2002). Lessons from the foot and mouth disease outbreak in the Netherlands in 2001. Rev. Sci. Tech. OIE 21, 711-721

R development core team (2011) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/

Willeberg, P., Halasa, T., AlKhamis, M., Boklund, A., Perez, A. and Enøe, C. (2012). Decisions on FMD vaccination informed by outbreak data and model predictions. Poster presentation, in the 13th ISVEE conference, Maastricht, the Netherlands, poster topic 08, poster number 34

World Organisation for Animal Health (OIE) (2011). Terrestrial Animal Health Code, Article 8.5.9

COMPARATIVE ASSESSMENT OF ECOLOGICAL FACTORS ASSOCIATED WITH REPORTING OF FOOT AND MOUTH DISEASE OUTBREAKS IN ENDEMIC REGIONS

F.O. MARDONES, R. HIJMANS, C. LOCKHART, J. PINTO AND A.M. PEREZ*

SUMMARY

Foot and mouth disease (FMD) is a highly infectious viral disease of cloven-hoofed animals that is endemic to certain parts of Asia, Africa and South America. Reporting FMD outbreaks to international organisations is mandatory and a number of repository disease data are available at local and global scales. The objective of this study was to compare the association between ecological factors and FMD outbreak reporting in the three continents in which the disease is still endemic. Geo-referenced FMD outbreak reports gathered from official sources (Empres-FAO, OIE, etc.) using the UC Davis Disease BioPortal system were collected for the last five years (2007-2012). A set of 28 environmental predictors was selected to summarise aspects of topography, vegetation, climate, animal and human's distributions. Maximum entropy (Maxent) was used to quantify the association between environmental predictors and FMD outbreak reporting for each continent, as indicated by the percentage of contribution of each predictor to the prediction (CP). Model fitness was assessed by computing area under the curve (AUC) values for predictions and values of AUC ≥ 0.9 were considered to demonstrate good model fitness. The most important predictor for FMD outbreak reporting was number of cattle for both Africa and Asia (CP = 29.3% and 20.8%, respectively). However, FMD outbreak reporting in South America was mostly predicted by anthropogenic biome (CP = 54.1%). Results are consistent with disparate epidemiological and ecological dynamics of the disease in South America, compared to Africa and Asia.

INTRODUCTION

Foot and mouth disease (FMD) is a highly infectious disease of cloven-hoofed animals caused by a virus (FMDV) that is still endemic in certain regions of Asia, Africa and South America (Saraiva, 2004; Kitching, 2005). Reporting FMD outbreaks to international organisations is mandatory because of the substantial economic impact that the disease imposes on affected countries.

---

* Andres M. Perez, Center for Animal Disease Modeling and Surveillance, University of California in Davis, One Shields Avenue, Davis, CA 95616, USA.
Email: amperez@ucdavis.edu

Identification of FMD outbreaks is typically based on the visualisation of FMD-like clinical signs and lesions and, consequently, reporting is mostly associated with the clinical identification of the disease. Because clinical manifestation and identification of the disease depends on a variety of ecological factors and because management and environmental conditions differ among FMD endemic settings, one would expect that the nature and extent of the association between FMD outbreak reporting and ecological factors would vary among FMD-endemic settings.

The aim of this study was to quantify and compare the power of ecological factors for predicting FMD outbreak reporting in Africa, Asia and South America. Results will contribute to identify factors and forces that rule the reporting process at regional scale and, ultimately, the disease epidemiology in regions in which FMD is yet to be eradicated.

MATERIALS AND METHODS

The maximum entropy method (Maxent) was used to quantify the association between ecological factors and FMD outbreak reporting. Briefly, Maxent is a machine-learning technique that characterises probability distributions from incomplete information (e.g., outbreak data) (Phillips et al., 2004) by finding the probability distribution of maximum entropy (the distribution that is most spread out or closest to uniform) subject to constraints imposed by the information available regarding the observed distribution of the outbreaks and environmental conditions across the study regions (Phillips et al., 2006). Outbreak information was evenly divided into training (for model building) and test (for model evaluation) datasets. Maxent works in an iterative process of rule selection, evaluation, testing, and incorporation or rejection. A method was iteratively (and until convergence) chosen from a set of possibilities (e.g., logistic regression, bioclimatic rules) to develop a rule (Pearson, 2007) with the highest predictive accuracy between iterations, which was incorporated into the model.

Results for the South American model were taken from another paper, in which the methodological procedures are described in detail (Mardones et al., unpublished), whereas the model for Asia and Africa and inter-region comparisons are original contributions of the paper here. FMD outbreak data for 2007-2012 were collected using the Disease Bioportal (http://fmdbioportal.ucdavis.edu/) system (Perez et al., 2009) from publicly available official databases, namely, the World Animal Health Information System (WAHIS), and the Global Animal Disease Information System (EMPRES-i) from the World Animal Health Organization (OIE) and Food and Agriculture Organization (FAO), respectively. Predictors included:

a) 19 climate-related covariates (Hijmans et al., 2005) from the WorldClim data set (http://www.bioclim.com): annual mean temperature, mean diurnal temperature range, isothermality, temperature seasonality, maximum temperature of warmest month, minimum temperature of coldest month, temperature annual range, mean temperature of wettest quarter, mean temperature of driest quarter, mean temperature of warmest quarter, mean temperature of coldest quarter, annual precipitation, precipitation of wettest month, precipitation of driest month, precipitation seasonality, precipitation of wettest quarter, precipitation of driest quarter, precipitation of warmest quarter, and precipitation of coldest quarter.

b) density distribution for cattle, sheep, goats, pigs and buffaloes, altitude for the topography component, human densities (anthropogenic biome) and croplands and grasslands that were obtained from the agriculture division at the FAO (http://www.fao.org/geonetwork/). The anthropogenic biome characterises the terrestrial biomes based on empirical analysis of global population, land use and land cover, resulting in eighteen categories (Ellis & Ramankutty, 2009).

The association between predictors and outbreak reporting was computed for the training data as the contribution of each predictor to the prediction (CP). Model fitness was evaluated by computing the area under the curve (AUC) of the receiver operating characteristics (ROC) curves, which was computed for the test dataset.

Analyses were carried out separately for each continent (Africa, Asia, South America) and using the packages raster for manipulating raster data, and dismo for Maxent modelling approaches implemented in the R software (R Development Core Team, 2012).

RESULTS

The model for Africa indicated an AUC value of 0.92, where predicted FMDV was associated with 14 environmental variables. Relative contributions of the environmental variables to the Maxent model ranged from 0.6% to 29.3%, the latter maximum provided by cattle density (Fig. 1).

In Asia, the predicted FMDV suitability distribution was associated with 18 environmental variables that most contributed to the fitting of the final model, and an AUC equal to 0.91. As in Africa's prediction, cattle densities contributed importantly (20.8%) to the final predicted model.

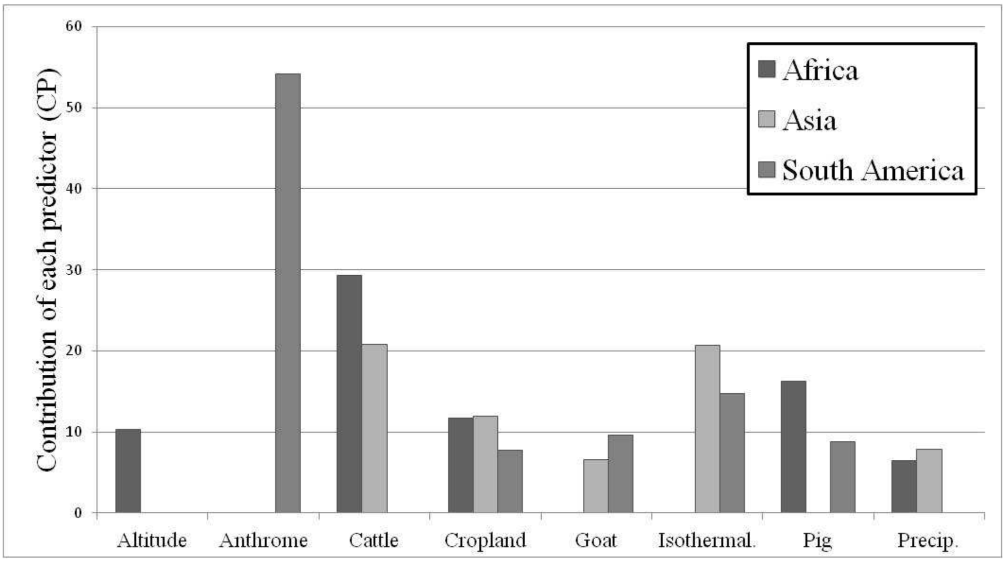


Fig. 1 The most important variables and their corresponding percentage of contribution retained at the final maximum entropy model fitted for each continent. [Isothermal. = Isothermality; Precip. = Precipitation]

Finally, the Maxent model for South America provided the maximum AUC value (AUC = 0.99), where the predicted value of the of the AB anthropogenic biome contributed with 54.1% to the final predicted model, in addition to 12 environmental variables that contributed in the range from 0.1% to 14.7% (Mardones et al, unpublished).

DISCUSSION

Results here provide quantitative estimates of the association between ecological variables and FMD outbreak reporting in the three continents in which FMD is still endemic. Interestingly, cattle density was the most important predictor for FMD outbreak reporting in Asia and Africa, but not for South America, where FMD was mostly predicted by the anthropogenic biome (Mardones et al, unpublished). These results are likely the consequence of disparate epidemiological disease and productive dynamics in Asia and Africa, compared to South America.

South America has made notable advances in the control of FMD and much of its territory is nowadays FMD free (either with or without vaccination) with sporadic FMDV incursions. Additionally, routine FMD mass vaccination is practised in much of the continent. Regionalisation is an important component of the control program with FMD outbreaks associated, mostly, with animal movements. Consequently, FMD outbreaks were predicted, with most, by the anthropogenic biome, which may serve as a proxy for markets and higher chances of clinical inspection by the veterinary services, compared to isolated regions (Mardones et al, unpublished).

Conversely, FMD outbreaks are frequent in large areas of Africa and Asia and coverage and effectiveness of vaccination campaigns are heterogeneous. Clinical signs are more evident in cattle than in other susceptible species, such as goats or sheep, and consequently, outbreak reporting was mostly predicted by cattle density in those regions.

In summary, results here provide metrics of the disparate influence that ecological factors have on FMD outbreak reporting in continents in which the disease is yet to be eradicated. These results are also consistent with lesser progress in the implementation of control and prevention programs for FMD in Africa and Asia compared to South America.

ACKNOWLEDGEMENTS

FOM is a PhD student supported by the Becas-Chile program of the National Commission for Scientific and Technological Research (CONICYT). This research was supported in part by the FAO and EpiX Analytics. No other conflict of interest exists.

REFERENCES

Ellis, E.C. and Ramankutty N. (2008). Putting people in the map: anthropogenic biomes of the world. Front. Ecol. Environ. 6, 439-447

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. and Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25, 1965-1978

Kitching, R.P. (2005). Global epidemiology and prospects for control of foot-and-mouth disease. Curr. Top. Microbiol. Immunol. 288, 133-148

Pearson, R.G. (2007). Species' distribution modeling for conservation educators and practitioners. Synthesis (American Museum of Natural History)

Perez, A.M., Zeng, D., Tseng, C.J., Chen, H., Whedbee, Z., Paton, D. and Thurmond, M.C. (2009). A web-based system for near real-time surveillance and space-time cluster analysis of foot-and-mouth disease and other animal diseases. Prev. Vet. Med. 91, 39-45

Phillips, S.J., Anderson, R.P. and Schapire, R.E. (2006). Maximum entropy modeling of species geographic distributions. Ecol. Modell. 190, 231-259

Phillips, S.J., Dudik, M. and Schapire, R.E. (2004). A maximum entropy approach to species distribution modeling. In: Proceedings of the Twenty-First International Conference on Machine Learning, Banff, Alberta, Canada, pp. 655-662

R Development Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. <http://www.R-project.org/>

Saraiva, V. (2004). Foot-and-mouth disease in the Americas: epidemiology and ecologic changes affecting distribution. Ann. N. Y. Acad. Sci. 1026, 73-78

DISEASE TRANSMISSION/CONTROL

SPECIFIC CATTLE EXCHANGES PRACTICES AND VECTOR-BASED TRANSMISSION SHAPE RIFT VALLEY FEVER VIRUS CIRCULATION IN MADAGASCAR HIGHLANDS

G. NICOLAS[*], B. DURAND, R. DUBOZ, R. RAKOTONDRAVAO AND V. CHEVALIER

SUMMARY

Last reported in 2008 in Madagascar, Rift Valley fever virus (RVFV) had a major impact on livestock. A previous study suggested that the trading network could partly support disease transmission in the highlands. The study reported here describes the cattle exchanges networks in a pilot area of Madagascar highlands and assesses their potential role in the diffusion of RVFV. Two trading practices were distinguished: usual trade and a traditional barter practice named *kapsile*. Social network analysis methods and logistic regression were used to analyse data. Networks topology suggested that preferential attachment mechanisms could play a role in exchange networks formation. Logistic regression showed that barter application was linked to RVFV circulation between villages whereas this was not the case for usual trade, which could rather support virus introduction from other parts of Madagascar.

INTRODUCTION

Rift Valley fever virus (RVFV) is a mosquito-borne zoonosis, belonging to the *Bunyaviridae* family and from the *Phlebovirus* genus. RVFV is known to cause high mortality levels in young ruminants and abortion storms in pregnant females. Infection in humans is generally associated with moderate influenza-like illness, but in 3% of cases severe ocular or neurological complications with sequelae or death may be observed. Transmitted between ruminants by mosquito bites, the virus may also be transmitted by direct contact with body fluids of viraemic animals (Gerdes, 2004). Human transmission may be vector-borne, or occur by direct contact with blood, tissues or body fluids from viraemic animals. Infection in humans is thus usually associated with an occupational exposure (Archer et al., 2011). Vector-borne transmission of RVFV is mainly supported by mosquitoes of the *Aedes*, *Culex* and *Anopheles* genera (Pépin et al., 2010).

The virus was first identified in 1930 in the Rift Valley in Kenya (Daubney et al., 1931). In 2000, RVFV circulation was reported in the Arabian peninsula (Ahmad, 2000). Since then, numerous outbreaks have been described in various eco-climatic regions of Africa. In early 2007, a large outbreak occurred across the Horn of Africa, first in Kenya (Sang et al., 2010),

---

[*] Gaëlle Nicolas, CIRAD – AGIRs Unit – Montpellier, France & ANSES – EPI Unit – Maisons-Alfort, France. Email: gaelle.nicolas@cirad.fr

Tanzania and Somalia (WHO, 2007) then in Sudan (Adam et al., 2010). A few months later, autochthonous RVF cases were reported in islands of the Indian Ocean: Comoros and Mayotte (Sissoko et al., 2009). The virus was probably introduced by the trade of live ruminants imported from eastern Africa during the previous epidemics of early 2007 (Paweska et al., 2010; Cêtre-Sossah et al., 2012).

RVFV was first isolated in Madagascar in 1979 from a pool of mosquitoes from the highlands. No human or livestock cases were reported that year (Morvan et al., 1991). Human and animal RVF cases were first reported in 1990-1991, during two consecutives rainy seasons (Morvan et al., 1992). In 2008, a large outbreak occurred, most of the regions being affected. Fifty nine human cases were confirmed in the highlands (of which 19 were fatal), despite the elevation and the low winter temperatures not being favourable to vectors (Andriamandimby et al., 2010). A phylogenetic study of the RVFV strains isolated in 1979, 1990-1991 and 2008 in Madagascar revealed that the strains involved in each outbreak were genetically different (Carroll et al., 2011) and probably introduced prior to each outbreak by cattle exchanges with the African mainland. The Anjozorobe district located in this highlands area was severely affected during the 2008 outbreak (Andriamandimby et al., 2010). A commune of this district, named Ambongamarina, was chosen as a pilot area to analyse RVFV circulation mechanisms in Madagascar highlands. A survey conducted in 2009 in individually identified bovines from this pilot area showed an IgG seroprevalence of 28% and suggested a recurrent circulation of the RVFV. This study also suggested (i) an absence of a wild rodent reservoir, (ii) a substantial mosquito-borne transmission, and (iii) a possible contribution of the cattle trade to the introduction and the persistence of the virus in this area (Chevalier et al., 2011).

Veterinary epidemiology usually uses social network analyses (SNA) to assess the involvement of animal movements in spreading diseases (Newman, 2002). In this case, animals are considered as a vehicle for pathogens which can spread between villages or regions and infect new animals during their contacts with the carrier. It is also possible to consider the contact network between actors as a medium for the spread rather than the livestock movements themselves. In some societies, traditional exchange practices may consist of behaviours that induce an exposure of cattle herds to disease transmission. In our study area, a traditional exchange practice named *kapsile* was identified in addition to usual trade. It consists of an animal barter which induces symmetrical movements of cattle: a breeder who barters cattle through *kapsile* always receives other cattle in return. However, to conclude the barter, the applicant has to exhibit his cattle around work areas. This leads to frequent contacts between the applicant herd and those of other breeders.

Two facets can thus be considered in the exchange network of our study area: (i) the movements of animals between herds/villages and (ii) the contacts between animals of herds/villages, induced by trade habits. The goal of this study was to describe the cattle exchanges practices in the pilot area, and to analyse the respective roles of these two facets in RVFV circulation in Madagascar highlands.

MATERIALS AND METHODS

Study area

Location: The study area was located 80 km to the north of the capital Antananarivo within the Anjozorobe district. This district is composed of 18 communes, among which our pilot area. This area is enclosed due to the primary rainforest corridor that surrounds it. The main route for the connection with the neighbouring communes is a track coming from the Anjozorobe city (in the west) and joining the Alaotra lake through the rainforest (in the north-east) (Fig. 1). Hard to join in the rainy season, people use carts pulled by cattle, bikes or walk to move in this area. It is made up of many rice and crop fields irrigated by a variety of different water systems constituting a suitable habitat for mosquitoes. Due to the high altitude of the area (about 1280m above sea level), the climate is characterised by two seasons: a dry and cold season from May to October, and a wet and warm season from November to April. The average winter temperature is about 9°C in August (Goodman et al., 2007), which makes this period unfavourable to mosquitoes and thus to vector-borne transmission.

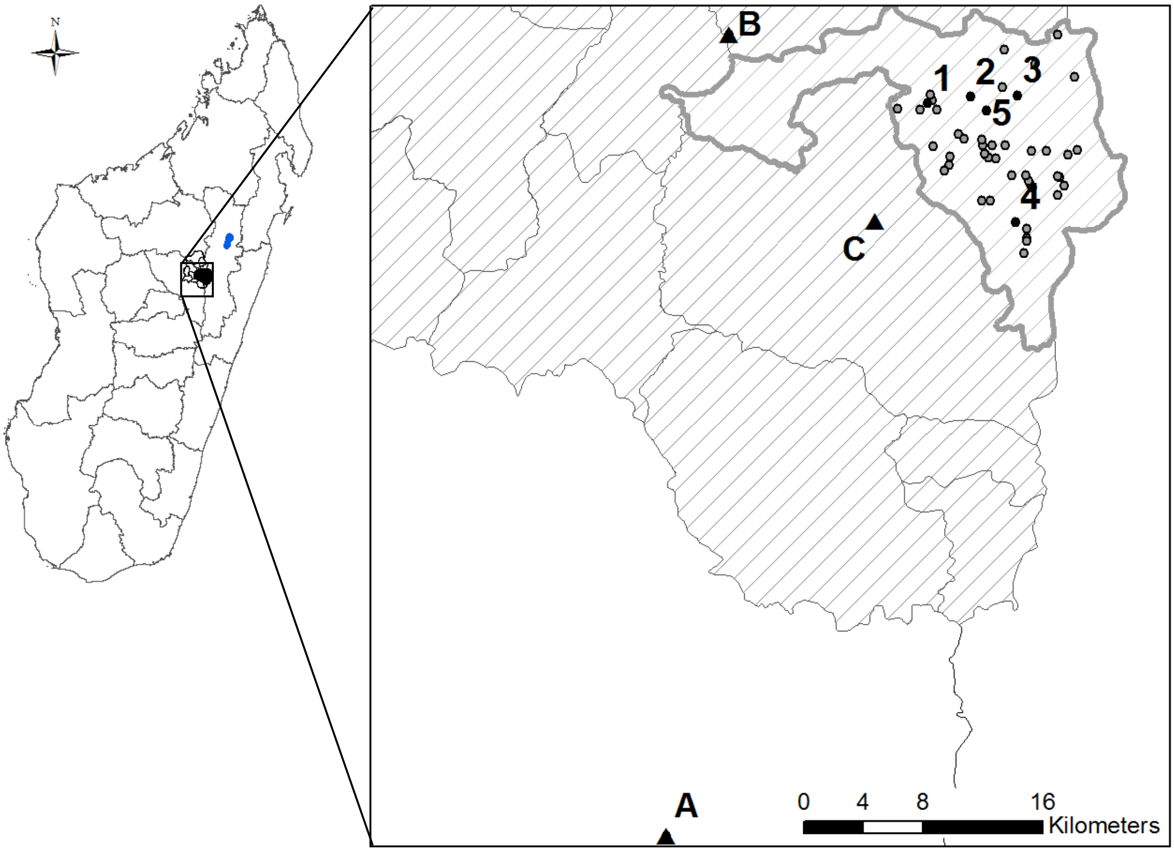


Fig. 1 Location of Anjozorobe village (C) and cattle markets (A: Ambatomena; B: Antanandava) that supply the pilot area (thick line) and the studied villages (dots) (1 to 4: higher centrality parameters; 5: market of Ambatolampy) of the Anjozorobe district (hatched)

In 2011, the cattle population of the pilot area was estimated by the Malagasy Veterinary Services at 2200 heads for the 52 villages. Herd size ranges from 1 to 15 cattle per household, which are usually employed for agricultural works (ploughing, planting or crop transportation), or slaughtered for traditional feasts, weddings, funerals, and for the national day. Cattle strongly contribute to food security in Madagascar (Minten & Barrett, 2008). Agricultural works are strongly seasonal, especially for rice. Rice planting is performed from November to January, followed by the harvest from March to May. The lucky farmers who possess rice field near permanent water points can benefit from a first period of rice planting in July-August which results in an additional harvest from December to January. This increased activity during agricultural work periods results in an increased need for labour and traction and, therefore, in suitable cattle for work. Cattle exchanges are thus also seasonal even if they may occur all year long.

Trading practices: The *usual cattle trade* involving a buyer and a seller with a monetary transaction between them was observed. After the transaction is concluded, the buyer goes to the village of the seller to bring the purchased cattle back home (the buyer village). Cattle were mainly bought from May to September and sold from December to March.

The *specific traditional trading habit* locally known as *"kapsile"* is mainly practiced by farmers during agricultural work periods (rainy season). This traditional practice allows the applicant to obtain animals suitable for agricultural works, despite the lack of money at that period of the year (before or during harvesting). Additionally, when money from the previous rice harvest is missing, some farmers may barter one valuable head of cattle against one of lower value in addition to money. This practice was termed below barter.

Cattle exchanges with outside locations were mainly formed by trade or barter with breeders of neighbouring communes, and trade at two identified markets (Fig. 1). The closest one, so-called Antanandava market, is located in the Anjozorobe district but 15 km from the study area. It supplies the commune with a local cattle breed (zebu) from a breeding area of the north of Madagascar (Betsiboaka region, mainly Andriamena and Tsaratanana district, respectively 100 and 140 km from the pilot area) which was heavily affected by RVFV in 2008 (Andriamandimby et al., 2010). According to the dealers, it takes about a week for cattle from the Betsiboaka region to join the Antanandava market. This consistent travel duration could thus support RVFV introduction in the study area. The farthest one, Ambatomena market, is located in the Manjakandrina district, along the road between the Anjozorobe city and the capital of Antananarivo. It is supplied with mixed breed cattle from farms and markets of the highlands.

Data collection

Networks data: A questionnaire survey involving breeders of the pilot area was carried out from February to July 2009. The average distance of the village to water was estimated from the average of the night pen distances to the nearest water point. The type of exchanges (barter vs trade), frequencies and periods of activity, number of cattle involved in each transaction, their breed, uses and the village origin/destination of the exchanged cattle were recorded. In the case of barter, the village of the applicant and of the receiver were recorded. Data stored in official registers of the commune were also collected: trade locations for breeders, monthly average number of cattle bought, trade periodicity, etc. The two markets located outside the pilot area but where breeders may directly have bought their cattle were investigated.

Serological data: Eight hundred and ninety four cattle from 43 villages were randomly selected in 2009, ear-tagged and blood sampled. Data on their age, sex and breed were also recorded. Samples were tested for anti-RVFV immunoglobulin (Ig) M and IgG as described in Chevalier et al. (2011) by the Madagascar Pasteur Institute. A seroprevalence of 28% was observed in 2009 (n=642) (Chevalier et al., 2011). Four hundred and eighty four cattle that tested seronegative in 2009 were sampled again in 2010 at their village of origin (n=39 villages) with an estimated incidence of 7% (95% confidence interval 5-10%), confirming that the virus was still circulating after the 2008 outbreak, with an intra-village seroconversion rate ranging from 0 to 20%. Thirty two were sold and sampled in another village of the study area. These cattle were still seronegative in 2010. One hundred and fifty eight animals were lost to follow-up, corresponding to a cattle replacement rate of 25% for 2009-2010.

Data analysis

Social network analysis methods (Wasserman & Faust, 1994) were used to describe and analyse the networks using the “sna” package (Butts, 2008) of R software (R Development Core Team, 2011). As link directions were known, three networks with directional relations were computed. Nodes were villages and ties were any movements of cattle connecting villages. Two networks were generated according to the type of exchange, trade or barter, termed below “trade cattle flow network” and “barter cattle flow network”, respectively. The third network included all the animal exchanges, termed below “full cattle flow network”. A fourth network was built in which a directed link from village A to village B was defined if a breeder of village A had been the applicant in a barter accepted by a breeder of village B. This network was called “barter application network”. Cattle age distributions by trading exchange modality (bought, sold or bartered cattle) and use (work or slaughtered), were compared using Student’s t-test.

Networks parameters: To describe each network, centrality parameters (degree, betweenness and random-walk betweenness), clustering coefficient and density were first calculated. In accordance with literature (Wasserman & Faust, 1994; Martínez-López et al., 2009), the degree was defined as the number of links leaving (outdegree) or arriving (indegree) to the node as well as the sum of both (freeman degree). The betweenness of a node has been expressed as the proportion of shortest paths between node pairs that pass through that node. Adapted from the betweenness, the random-walk betweenness considers any path between nodes, not just the shortest (although it still gives more weight to short paths). It was the proportion of random walks between node pairs that pass through that node. The network density was the proportion of observed ties among the possible ties between nodes. The clustering coefficient is a measure of the degree to which nodes in a network tend to cluster together (Newman, 2003). Network-level clustering coefficient, degree, betweenness and random-walk betweenness correspond to the average of the node-level values of the corresponding indicators.

Network structure and occurrence of seroconversion: Logistic regression was used to analyse RVFV circulation in the study area. The outcome variable was the occurrence of seroconversion in the village. The explanatory variables were the node-level network descriptive parameters for the barter application and trade cattle flow networks. The following factors, potentially linked to the occurrence of seroconversion, were also included as predictors in the model: the 2009 seroprevalence (immunity level), the mean distance from the village to the nearest water point (Chevalier et al., 2011) and the occurrence of links from markets and from villages located outside the study area (binary variable: 0 if links were absent, 1 otherwise). Eight other models (Table 2) were tested in the same way but using centrality parameters of different networks. The nine models were compared using the Akaike Information Criterion (AIC).

Absence of significant collinearity between predictors was assessed by verifying that the Variance Inflation Factor (VIF) was lower than 5. The numbers of animals that were seronegative in 2009 and sampled in 2010 were used as weights in logistic regression. The node-level centrality parameters are by definition non-independent. A bootstrap procedure (Pesarin, 2001; Sinha, 2009) was implemented to assess the significance of model coefficients. Random permutations of the original adjacency matrix were generated to create 10,000 new networks. The logistic model was fitted for each of these. For each model coefficient, the empirical p-value was the proportion of the 10,000 networks for which the value was higher (or lower) than that obtained with the real network. Odds ratios (OR) were calculated for the model that had the lowest AIC, using the interquartile range of the distributions of quantitative variables.

RESULTS

The survey involved 48 villages and 386 breeders. Only one “little market”, so-called Ambatolampy market, was found inside the pilot area (Fig. 1). Cattle sold there were mainly unsold cattle from the Ambatomena market, located about 60 km from the pilot area.

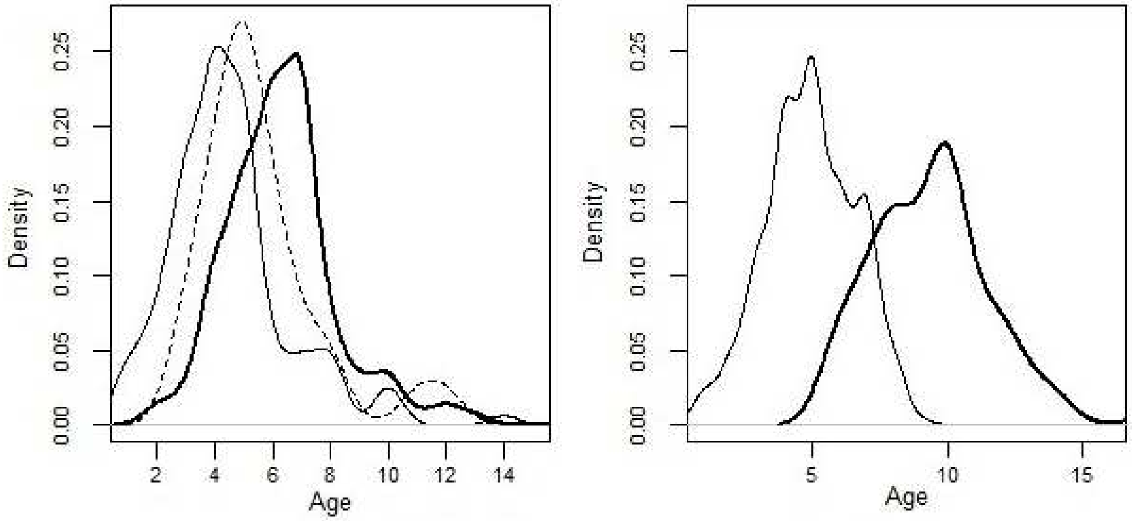


Fig. 2 Age distribution of cattle exchanged in 2010 in the study area according to the type of exchange (left, dashed line: barter; thick line: cattle sold; thin line: cattle bought), and to the destination of the animals (right, work: thin line, slaughter: thick line)

In one year, among the 40% of breeders that bought cattle (156/386), 71% purchased directly from a breeder whereas the other 29% bought on a market. In the latter case, 33% of cattle (n=45) came from Antanandava and 67% from Ambatomena. Concerning the purchase, 22% of buyers (n=156) had bought cattle more than once a year (vs 40% of barter users) whereas 78% did so to replace an animal too old to work efficiently (vs 60% barter users). Among cattle exchanges reported by breeders, 77 barters were registered and motivation of 30 of them was known (3 for slaughter, 27 for work); 126 cattle were bought (6 for slaughter and 120 for work) and 128 were sold (23 for slaughter and 105 for work). During the study period, the average age of cattle involved in barter and trade was approximately 5.5 years. The average age of bought cattle (4.5 years) was nearly 1 year younger than the average age of sold cattle (6.4 years). Moreover, bought cattle were significantly younger (p=0.02) than bartered cattle, themselves significantly younger than sold cattle (p=0.05). Cattle exchanged for slaughter were on average 4.5 years older than cattle exchanged for work (p<0.0001) (Fig. 2).

The full cattle flow network was composed of 168 arcs. Three isolated nodes were identified and correspond to villages without any reported cattle exchange with the other villages of the study area. Network-level centrality parameters and clustering coefficients were low (Table 1). The density value of the full cattle flow network showed that 7% of possible node pairs were actually linked by cattle exchange. The three cattle flow networks (full, trade and barter) contained a single component in which the average length of the shortest path between node pairs was approximately of 3 links, the maximal value (diameter) being 6 links. On average, a given village was directly connected with 7 other villages by cattle movements (average degree), was located on 2.8% of the shortest paths (average betweenness) and on 5% of any path (average random path betweenness) between node pairs. The trade cattle flow networks appeared to have lower centrality parameter values than those of barter networks (Table 1). However, a low but significant association was observed between them (0.22, p<0.0001), concerning 24 links between 18 nodes. In both of them (trade and barter cattle flow), nodes with the highest values for centrality parameters (respectively 2 and 3, Fig. 1) were different suggesting that villages favoured by users of both exchange habits may be different, and were located in the vicinity of the Ambatolampy market. This small market had low values of degree and betweenness and was not a major place for cattle exchange. Both networks were heterogeneous meaning that the majority of nodes were weakly linked, whereas a few nodes were linked with many other nodes.

Table 1. General parameters of the full, trade, and barter cattle flow networks

Parameters	Full cattle flow network	Trade cattle flow network	Barter cattle flow network
Number of nodes	48	48	48
Number of links	168	88	104
Diameter	6	6	6
Clustering coefficient	0.31	0.22	0.28
Average path length	2.7	3.0	2.7
Density	0.07	0.04	0.05
Average degree	7	3.7	4.3
Average betweenness	61.3	41.2	33.6
Normalized average betweenness	0.028	0.019	0.016
Random-walk betweenness	0.051	0.018	0.029

Seroconversions were observed in 26 of 39 villages. The logistic model with the lowest AIC was model 5 (Table 2) which combined centrality parameters (degrees) of the trade network and of the barter application network. This model had an AIC difference of 41 with the next best model (model 4) which considered the barter cattle flow network without distinction in the practice (applicant or receiver). Model 5 showed a protective effect of cattle introduction by trade (OR=0.3, p=0.03), and of the village distance to the closest water point (p=0.002). The OR corresponding to an increase of 160 meters to the nearest water point (inter-quartile range) was 0.2. Conversely, the presence of a barter applicant in the village was a seroconversion risk factor (OR=8, p=0.007). No significant association was observed between the occurrence of seroconversion and the level of herd immunity in 2009, nor with cattle introduction from foreign origin.

Table 2. Studied logistic models of the occurrence of seroconversions in the villages of the study area between 2009 and 2010, and corresponding Akaike Information Criterion (AIC) values

Model	Logistic Regression	AIC
1	SC ~ degree + dist.water + seroprev.09 + Foreign.origin	317
2	SC ~ indeg + outdeg + dist.water + seroprev.09 + Foreign.origin	318
3	SC ~ degree$_T$ + degree$_K$ + dist.water + seroprev.09 + Foreign.origin	307
4	SC ~ indeg$_T$ + outdeg$_T$ + degree$_K$ + dist.water + seroprev.09 + Foreign.origin	292
5	SC ~ indeg$_T$ + outdeg$_T$ + indeg$_A$ + outdeg$_{A +}$ dist.water + seroprev.09 + Foreign.origin	251
6	SC ~ betweenness + dist. water + seroprev.09 + Foreign.origin	313
7	SC ~ betweenness$_T$ + betweenness$_K$ + dist.water + seroprev.09 + Foreign.origin	312
8	SC ~ rbtw + dist.water + seroprev.09 + Foreign.origin	343
9	SC ~ rbtw$_T$ + rbtw$_K$ + dist. water + seroprev.09 + Foreign.origin	317

SC: occurrence of seroconversions between 2009 and 2010; indeg = indegree; outdeg = outdegree; rbtw = random-walk betweenness; dist.water = village distance to the nearest water point; seroprev.09 = seroprevalence in 2009; Foreign.origin = trade on markets or villages located outside the study area. Subscripts T, K and A refer to the trade cattle flow network, barter cattle network and barter application network, respectively.

DISCUSSION

The high incidence rate calculated in 2009-2010 shows that the virus kept circulating in the Anjozorobe area after the 2008 outbreaks. Despite an average seroconversion rate of 7% in the pilot area, no clinical case has been reported, confirming that RVFV may circulate silently but intensively.

Cattle of the area are used for agricultural work and are approximately replaced every 4 years. Cattle were bought and bartered for work whereas they were also sold for slaughter. The age of exchanged cattle varied according to the exchange practice (barter and trade), as well as according to the purpose of the exchange (slaughter or work). The distribution of node degrees and of local clustering coefficients is consistent with a preferential attachment in the nodes matching of the barter and trade cattle flow networks (Barabàsi et al., 1999). A low correlation between the barter and the trade cattle flow networks was observed. Distance between villages did not seem to influence the establishment of links. The study highlighted that villages located close to each other were not necessarily connected by cattle exchanges. Moreover, when choosing a supplier, long distances do not seem to deter breeders of this area. Some surveyed breeders bought cattle on the market of Ambatomena even though it is located about 60 km from the pilot area. The trade link would thus be established due to the better reputation of some breeders or villages. Barter provides animals used for work without any additional cost or conversely with earning money but without reducing the herd size.

Trade and barter are very different exchanges practices. Their careful description is necessary for the understanding of RVFV circulation mechanisms. Most breeders exchanged cattle occasionally according to their needs, and few of them are actually breeders (i.e. keep animals for reproduction) or full-time cattle traders. However, specific breeders or villages are known for their high quality in breeding or trade (personal observations). When cattle are bought, no contacts between animals of both villages occur. Bought cattle are either young and used for work and breeding, or old and slaughtered for meat consumption. Conversely, barter applicants would show their cattle to at least a few breeders sometimes living in several villages before concluding the deal. This implies many contacts between the herd of the barter applicant and those of other breeders.

A statistical association between network centrality parameters and the occurrence of seroconversion in villages of the pilot area was found but opposite effects were observed. Considering trade, buying cattle from different villages (indeg_T) had a protective effect against RVFV infection. Conversely, barter applicants (outdeg_A) appeared to be risk factors for their own village. This could be explained by age of the introduced animals and the exchange practices themselves. Chevalier et al. (2011) did not find seropositive cattle younger than 3 years, whereas the proportion of seropositive animals increased with age. This study showed that bought cattle were either old (often immune) or young. Old cattle were immediately slaughtered whereas young animals, too young to work immediately, were kept at pasture and reared for future work without going into rice fields which are larval habitats for the potential mosquito-vectors. Moreover, most of these young cattle were bought during the winter when the mosquitoes are absent. Exposure of these young animals to RVFV is thus probably limited. Conversely, when animals are bartered through *kapsile,* it occurs during the rainy season and during periods of vector activity (at the end of the day), near larval habitats for the competent mosquito species (flooded rice fields). Herds of barter applicants are thus more exposed to the mosquito bites. Furthermore, most of the time animals are bartered during work periods and directly needed in the rice fields where they are thus probably more exposed to RVFV than those introduced by trade. Increasing the outdegree in the barter application network is thus a risk factor for the village of the applicant.

Despite the hypothesis of RVFV introduction by trade coming from markets and/or villages located outside the study area, no statistical association with the occurrence of seroconversions was observed. Nevertheless, it could represent a real risk with a low occurrence probability that could not be revealed by this study. The expected protective effect of immunity cover was not observed. This could be explained by a lack of power due to a small amount of recorded data.

This study has highlighted original cattle exchange practices that induce exposure to RVFV and could be important factors in its spread. The low activity level of the communal market results from its paucity in cattle choice. This seems to be compensated by the breeders' high interest in barter and in larger markets. Breeders are prepared to travel long distances in order to choose the most suitable cattle, confirming the idea of the importance attached to the cattle choice by breeders. The involvement of the two facets of the trading exchanges practices (trade movement vs contact in barter) in RVFV circulation is related to the seasonality of work, to money availability and to vector proliferation. Cattle were mainly introduced after the rice collection, during the winter, when money is abundant but in absence of vectors. Conversely, barter through *kapsile* mainly occurred during the work period, during the rainy season and thus during the vector proliferation period, when the money becomes scarce. Due to links with cattle markets located in areas with favourable climatic condition for RVFV circulation and that were affected during the 2008-2009 outbreak, the trade network could support virus introduction from other parts of Madagascar. The negative relation of the distance to the nearest water point suggests a role of vectors in RVFV circulation in this area despite unfavourable climatic conditions. After virus introduction in a village, vector-based transmission could be responsible for within-village circulation. The barter application network could promote RVFV circulation between villages due to cattle exposure implied in this practice. A survey conducted in 2010 on the markets of the Madagascar highlands, as well as a third serological survey conducted in the study area, should bring more information on highlands cattle trade and on the epidemiological processes involved in the recurrent transmission of RVF in this temperate and mountainous area.

ACKNOWLEDGEMENTS

We acknowledge the personnel of the Veterinary Services and breeders of Madagascar for their involvement. This study was conducted by CIRAD and ANSES on the basis of data collected in the project entitled “Rift Valley fever in the Indian Ocean Islands” (RIFT-OI).

REFERENCES

Adam, A.A., Karsany, M.S. and Adam, I. (2010). Manifestations of severe Rift Valley fever in Sudan. Int. J. Infect. Dis 14, e179-180

Ahmad, K. (2000). More deaths from Rift Valley fever in Saudi Arabia and Yemen. Lancet 356, 1422

Andriamandimby, S.F., Randrianarivo-Solofoniaina, A.E., Jeanmaire, E.M., Ravololomanana, L., Razafimanantsoa, L.T., Rakotojoelinandrasana, T., Razainirina, J., Hoffmann, J., Ravalohery, J.P., Rafisandratantsoa, J.T., Rollin, P.E. and Reynes, J.M. (2010). Rift Valley fever during rainy seasons, Madagascar, 2008 and 2009. Emerg. Infect. Dis. 16, 963-970

Archer, B.N., Weyer, J., Paweska, J., Nkosi, D., Leman, P. A., Tint, K.S. and Blumberg, L. (2011). Outbreak of Rift Valley fever affecting veterinarians and farmers in South Africa, 2008. S. Afr. Med. J. 101, 263-266

Barabàsi, A.-L., Albert, R. and Jeong, H. (1999). Mean-field theory for scale-free random networks. Physica A 272, 173-187

Butts, C.T. (2008). A package for managing relational data in R. J. Stat. Softw. 24 (*version 2.2-0*)

Carroll, S.A., Reynes, J.M., Khristova, M.L., Andriamandimby, S.F., Rollin, P.E. and Nichol, S.T. (2011). Genetic evidence for Rift Valley fever outbreaks in Madagascar resulting from virus introductions from the east African mainland rather than enzootic maintenance. J. Virol. 85, 6162–6167

Cêtre-Sossah, C., Pédarrieu, A., Guis, H., Defernez, C., Bouloy, M., Favre, J., Girard, S., Cardinale, E. and Albina, E. (2012). Prevalence of Rift Valley fever among ruminants, Mayotte. Emerg. Infect. Dis. 18, 972-975

Chevalier, V., Rakotondrafara, T., Jourdan, M., Heraud, J.M., Andriamanivo, H.R., Durand, B., Rollin, P.E. and Rakotondravao, R. (2011). An unexpected recurrent transmission of Rift Valley fever virus in cattle in a temperate and mountainous area of Madagascar. PLoS Negl. Trop. Dis. 5, e1423

Daubney, R., Hudson, J. and Garnham, P.C. (1931). Enzootic hepatitis or Rift Valley fever: an undescribed disease of sheep, cattle and man from east Africa. J. Pathol. Bacteriol. 34, 545-579

Gerdes, G.H. (2004). Rift Valley fever. Rev. Sci. Tech. 23, 613-623

Goodman, S.M., Raselimanana, A.P. and Wilmé, L. (2007). Inventaires de la faune et de la flore du couloir forestier d'Anjozorobe-Angavo, Analamerana et Andavakoera. Ed. Sciences biologiques. Recherche pour le développement, Antananarivo, Madagascar. 216

Martínez-López, B., Perez, A.M. and Sánchez-Vizcaíno, J.M. (2009). Social Network Analysis. Review of general concepts and use in preventive veterinary medicine. Transbound. Emerg. Dis. 56, 109–120

Minten, B. and Barrett, C.B. (2008). Agriculture, pauvreté rurale et politiques économiques à Madagascar. Ed. Ilo Working Paper, Cornell University, Ithaca, New York

Morvan, J., Fontenille, D., Saluzzo, J.F. and Coulanges, P. (1991). Possible Rift Valley fever outbreak in man and cattle in Madagascar. Trans. R. Soc. Trop. Med. Hyg. 85, 108

Morvan, J., Rollin, P.E., Laventure, S., Rakotoarivony, I. and Roux, J. (1992). Rift Valley fever epizootic in the central highlands of Madagascar. Res. Virol. 143, 407-415

Newman, M.E.J. (2002). Spread of epidemic disease on networks. Phys. Rev. E. 66, 016128

Newman, M.E.J. (2003). The structure and function of complex networks. SIAM Review 45, 167-256

Paweska, J.T., Weyer, J., van Vuren, P.J., Kemp, A., le Roux, C., Leman, P., Grobbelaar, A., Archer, B., Thomas, J., Last, R., de Klerk, G., Pienaar, H., Blumberg, L. and Swanepoel, R. (2010). Update on the 2010 Rift Valley fever outbreak in South Africa. ArboZoonet News 5, 7-13. Available on: http://www.arbo-zoo.net/_data/arbozoonet-news_No15.pdf

Pépin, M., Bouloy, M., Bird, B.H., Kemp, A. and Paweska, J.T. (2010). Rift Valley fever virus (Bunyaviridae: Phlebovirus): an update on pathogenesis, molecular epidemiology, vectors, diagnostics and prevention. Vet. Res. 41, 61

Pesarin, F. (2001). Multivariate Permutation Test: with Applications in Biostatistics. Ed. John Wiley & sons, LTD, University of Padova, Italy

R Development Core Team (2011). R: A language and environment for statistical computing. Vienna, Austria. R Foundation for Statistical Computing. Available on: http://www.R-project.org

Sang, R., Kioko, E., Lutomiah, J., Warigia, M., Ochieng, C., O'Guinn, M., Lee, J.S., Koka, H., Godsey, M., Hoel, D., Hanafi, H., Miller, B., Schnabel, D., Breiman, R.F. and Richardson, J. (2010). Rift Valley fever virus epidemic in Kenya, 2006/2007: The entomologic investigations. Am. J. Trop. Med. Hyg. 83, 28–37

Sinha, S.K. (2009). Bootstrap test for variance components in generalized linear mixed model. Can. J. Stat. 37, 219-234

Sissoko, D., Giry, C., Gabrie, P., Tarantola, A., Pettinelli, F., Collet, L., D'Ortenzio, E., Renault, P. and Pierre, V. (2009). Rift Valley fever, Mayotte, 2007-2008. Emerg. Infect. Dis. 15, 568-570

Wasserman, S. and Faust, K. (1994). Social Network Analysis: Methods and application. Ed. Cambridge University Press, Cambridge, 825

WHO (2007). Outbreaks of Rift Valley fever in Kenya, Somalia, and United Republic of Tanzania, December 2006-April 2007. Weekly epidemiological record 82, 169-180

USE OF MULTILEVEL MODELS TO EVALUATE RISK FACTORS AT THE WILDLIFE-DOMESTIC INTERFACE FOR BOVINE TUBERCULOSIS IN SOUTH-CENTRAL SPAIN

B. MARTÍNEZ-LÓPEZ*, J.A. BARASONA, C. GORTÁZAR, V. RODRÍGUEZ-PRIETO,
J.M. SÁNCHEZ-VIZCAÍNO AND J. VICENTE

SUMMARY

Eradication of tuberculosis (TB) in cattle is a priority in the EU but seems a goal far from being accomplished. The identification of risk factors contributing to TB persistence is key to better understand and more cost-effectively eradicate this complex zoonotic disease worldwide. This study aimed to identify farm-level factors contributing to TB occurrence, new infection and persistence in one of the most TB-prevalent regions in South-Central Spain, Ciudad Real, for which detailed information on cattle and wild ungulates was available for up to 5 years (2007-2011). Data were analysed using multilevel logistic regression models. Results revealed a significant association between wild ungulate management, abundance and proximity and TB risk in cattle farms. Beef and bullfighting cattle farms, particularly those of large size and near to goat farms, were at highest TB risk. These results may support policies to better prevent/control TB in this and other similar territories.

INTRODUCTION

Tuberculosis (TB), caused by *Mycobacterium tuberculosis* complex (MTC), particularly *M. bovis* and *M. caprae*, is a zoonotic bacterial disease that affects a wide range of domestic and wild species all over the world (Alexander et al., 2002; Smith et al., 2011). Although TB eradication has been a major objective for developed countries in the last decades, this goal is still far from being accomplished (Reviriego Gordejo & Vermeersch, 2006; Schiller et al., 2011). Despite the tremendous efforts and investment to achieve TB eradication, some countries such as UK, Ireland, France or Germany have experienced an increase of TB cases in recent years (Schiller et al., 2011). In Spain, the overall TB trend is a small decrease in incidence, but some regions, like South-Central Spain, are still considered to have a high TB prevalence (>3%) despite the intensive and costly (34.7 million € during 2011) eradication programme in place (MAGRAMA, 2012). Moreover, in Mediterranean extensive farming habitats from Iberia (i.e. Spain and Portugal), there is a tendency of private livestock farms to

---

* Beatriz Martínez-López, IREC (CSIC-UCLM-JCCM), Ronda de Toledo s/n; 13005, Ciudad Real (Spain) and VISAVET, University Complutense of Madrid, Av. Puerta de Hierro s/n; 28040, Madrid (Spain). Email: beatriz@sanidadanimal.info; Beatriz.Martinez@uclm.es

become hunting estates, gathering wild and domestic ungulates in the same geographical area, which creates significant risks for disease transmission (Phillips et al., 2003).

In the last years, many risk factors for TB occurrence at an individual/animal, herd and regional/international level have been identified (see Humblet et al., 2009 for a review). However, there is still a lack of knowledge and controversy on (1) the risk factors that are mostly contributing to TB transmission under the different and, usually complex, epidemiological scenarios, (2) the specific role of different wildlife species in TB maintenance and transmission to domestic populations, and (3) the recommended and most cost-effective control and eradication measures that should be implemented in specific infected areas, given the apparent lack of effectiveness of some of the strategies currently in place (Reviriego Gordejo and Vermeersch, 2006).

Some studies in the United Kingdom and New Zealand described the association between movement of animals, herd size, management practices and/or badger presence and TB occurrence (Gilbert et al., 2005; Porphyre et al., 2008). Similarly, a recent Belgian study described previous (i.e. historical) occurrence of TB in the herd, proximity of an outbreak, cattle density and cattle movements as the main risk factors for TB (Humblet et al., 2010). Proximity and contact with wildlife species have also been hypothesised to play an important role in TB transmission (Humblet et al., 2010; Allepuz et al., 2011; Rodríguez-Prieto et al., 2012). This suggestion has been reinforced after finding the same MTC spoligotypes of cattle in some wildlife species (Aranaz et al., 1996, Gortazar et al., 2005).

The study presented here aimed to explore the potential risk factors for TB occurrence, new infection and persistence in cattle farms from one of the most problematic and TB-prevalent regions in South-Central Spain, referred to as Ciudad Real (CR). Of particular interest was the wildlife-domestic interface and evaluating the potential role of wild ungulates abundance, proximity and management practices in TB occurrence, new infection and/or persistence. In order to accomplish this goal, multilevel logistic regression models were used to evaluate all cattle farms in CR for which information regarding TB status, cattle movements, management practices and proximity/abundance to wildlife reservoirs was available. Methods and results presented here may help to reduce TB prevalence in CR and may support policies in other regions and countries with similar epidemiological conditions.

MATERIALS AND METHODS

Study area and data

The study was performed in CR (37º 13’ 48” N to 39º 31’ 43” N in latitude; 2º 25’ 54” W to 6º 34’ 06” W in longitude; 19,813 km2). Ciudad Real has a complex epidemiological situation, with coexistence of wild and domestic ungulates under very diverse management practices. It is also one of the provinces that have traditionally registered the highest levels of TB herd prevalence and incidence in Spain (Allepuz et al., 2011; MAGRAMA, 2012). For that reason, and because detailed information on both cattle and wildlife was available, it was selected for this study. In this region, red deer and wild boar (and, to a lesser extent, fallow deer) may exist as ‘managed’ populations, which stay on large fenced hunting estates and are kept at high densities, with supplementary food provided. In contrast, ‘unmanaged’ wild ungulate populations occur on open (i.e. no fenced) estates where they are completely free-living and not specifically managed for hunting, with no supplementary feeding provided.

The artificially high densities of the managed populations and the provision of additional food and water sources, which encourages wild ungulates to aggregate, are likely drivers of TB maintenance (Vicente et al., 2007). In unmanaged populations however, the aggregation of wild ungulates around food and water sources during periods of drought may pose a similar increased transmission risk (Vicente et al., 2007).

Data regarding farm size (i.e. number of cattle on the farm), type of production on the farm (i.e. beef, dairy or bullfighting cattle), species produced on the farm (only cattle or cattle with goats, sheep or pigs), location (x and y coordinate of the farm centroid), number of movements and animals leaving and entering the farm per year and TB status of the farm (i.e. prevalence, incidence, number of samples taken, number of positive animals) were available for each farm in the CR region from 2007 to 2011. Detailed information regarding proximity, abundance and management of wildlife species in nearby (<2 km) hunting estates was also available for each farm. Specifically, data from hunting estates were processed with ArcGIS 9.2 (ESRI) using the annual hunting bags (i.e. number of animals hunted per hunting event) for different species during the study period, and a layer including the perimeters of all hunting estates, which is the smallest management unit. Information on the demographics and TB status of cattle farms, annual hunting bag per estate and species, geographical borders and the fencing situation of each estate was provided by the Castile-La Mancha regional government. Environmental variables (land use obtained from the Corine Land Cover 2006) were also available for each farm.

Statistical analyses

Three multilevel logistic regression models were used to evaluate the potential association between hypothesised risk factors and TB occurrence (Model I), new infection (Model II) and persistence (Model III) on farms in the CR region. Models were expressed as follows:

$$y_{ij} \sim Binomial(n_{ij}, \pi_{ij})$$

$$logit(\pi_{ij}) = \log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_{0j} + \beta_1 X_{1ij} + \beta_2 X_{2ij} + \cdots \beta_k X_{kij} + \mu_j$$

where y_{ij} is the binary dependent variable for farm i at municipality j; π_{ij} is the expected probability of TB occurrence (Model I), new infection (Model II) or persistence (Model III), respectively; β_{0j} is the intercept; β_{1ij}, β_{2ij}… β_{kij} are the slopes; and μ_j is the random effect, which indicates the dispersion among farms from the same municipality and which follows a normal distribution of the form $N(0, \sigma_j^2)$. The dependent variable (y_{ij}) of Model I was whether the farm had been declared as TB positive or not (coded as 1/0, respectively) in any of the five years of study. A “TB-positive” farm was defined as a farm with at least one animal testing positive for TB using the intradermal tuberculin test (IDTT). In Model II, the response variable was whether or not the farm had become newly infected during the five years of study, defining a “new infection” as the observation of at least one positive animal (detected by the IDTT) after one year or more of absence of disease on farm. The response variable of Model III was coded as 1 if disease was diagnosed on the farm for two or more consecutive years, or 0 otherwise. For each of the three models the same predictors (up to 51) were used, representing farm demographics and management, cattle movements, proximity to other farms and historical persistence of TB on the farm, wildlife proximity, management and abundance and land use around the cattle farms. Quantitative variables were also transformed and tested as predictors in their binomial form (using the median as a cut-off point) in the three models.

All main effects and two-way interaction terms were evaluated as potential predictors into each of the three models. The selection of the best predictors for each of the three models was conducted by using the forward selection process, using AIC as criterion for model selection (Venables & Ripley, 2002). The significance level to enter a variable into the model was p <0.10. Collinearity was evaluated using the variance inflation factor (VIF) (Belsley et al., 1980). We used residual analysis (i.e evaluation of deviance residuals) and receiver operating characteristic (ROC) curves with area under the curve (AUC) values to evaluate the goodness of fit and the predictive ability of the model. All statistical analyses were conducted in R, v. 2.15.2 (R Development Core Team, 2012).

RESULTS

TB status and distribution in Ciudad Real

A total of 662 farms were included in the analysis, of which 588 (88.8%) were beef cattle farms, 57 (8.6%) were dairy cattle farms and 17 (2.6%) were bullfighting cattle farms. The median [95% confidence interval] number of cattle in beef, dairy and bullfighting farms was 76 [8, 412], 256 [64, 686] and 37 [3, 310], respectively. A total of 237 (35.8%) farms were found to be TB positive at least one year between 2007 and 2011, of which 171 (25.8%) were newly infected and 117 (17.6%) were considered to be persistently infected.

Risk factors for TB occurrence, new infection and persistence at farm level in Ciudad Real

Significant risk factors for TB occurrence, new infection and persistence were having an extensive production system on farm (i.e. beef/bullfighting), having a large farm size in terms of number of cattle and a high (> median) number of fenced big game estates within a 2-km radius around the farm (Table 1). High distance to sheep farms was also a significant risk factor for TB occurrence (OR=1.21) and new infection (OR=1.19). A high number of hunted deer per hunting estate within a 2-km radius around the farm (OR=1.20), a high minimum distance to open big game estates (OR=1.45) and a high average number of outgoing cattle shipments per year (OR=1.52) were significant risk factors for new TB infection. Historical (2003-2006) TB persistence (i.e. more than 1 year positive) on the farm was also a significant risk factor for TB occurrence (OR= 4.23) and persistence (OR=3.84).

A significant ($p < 0.05$) or borderline significant ($p < 0.10$) protective factor for TB occurrence (OR=0.71) and new infection (OR=0.64) was a high distance to goat farms (Table 1). High average number of cattle leaving the farm per year was also a significant protective factor for TB occurrence (OR= 0.79) and persistence (OR=0.71). Having a high number of open big game estates within a 2-km radius around the farm was also a significant protective factor for TB persistence (OR=0.62). There were no significant interaction terms. The three models had a good predictive ability, with AUC values of 0.824, 0.767 and 0.852, for Models I, II and III, respectively (Table 1).

Table 1. Odds ratios (OR) and 95% confidence intervals (CI) obtained for multilevel logistic regression models of TB occurrence (Model I), new infection (Model II) and persistence (Model III) on cattle farms in Ciudad Real

Variable name (description)	Model I OR [95% CI]	Model II OR [95% CI]	Model III OR [95% CI]
Calif (Type of production on farm. Coded as 1=beef and bullfighting; 0=dairy)	3.03 [1.18, 7,81][a]	3.01 [1.11, 8.13][a]	3.69 [0.85, 16.09][b]
Census (Standardised average number of animals on farm during the studied period)	1.99 [1.45, 2.72][a]	1.33 [1.10, 1.61][a]	2.19 [1.53, 3.13][a]
nHSM.C (Number of fenced big game estates within a 2-km radius around the farm. Coded as 1=above median, 0=below median)	1. 75 [1.15, 2.68][a]	1.55 [0.97, 2.49][b]	1.79 [1.06, 3.03][a]
nHSM.O (Standardised number of open big game estates in a 2-km buffer around the farm)	NA	NA	0.62 [0.46, 0.85][a]
Deer.HS (Standardised number of deer hunted per game estate within a 2-km radius around the farm)	NA	1.20 [0.97, 1.47][b]	NA
Dist.HSM.O (Minimum distance -in meters- to open big game estates. Coded as 1=above median, 0=below median)	NA	1.45 [0.97, 2.16][b]	NA
Dist.goat (Minimum distance –in meters- to goat farms. Coded as 1=above median, 0=below median)	0.71 [0.46, 1.09][b]	0.64 [0.42, 0.97][a]	NA
Dist.sheep (Standardised minimum distance -meters- to sheep farms)	1.21 [0.99, 1.46][b]	1.19 [0.99, 1.43][b]	NA
N.out-mov.cattle (Average number of outgoing cattle shipments per year. Coded as 1=above median, 0=below median)	NA	1.52 [1.01, 2.29][a]	NA
N.out-anim.cattle (Standardised average number of cattle leaving the farm per year)	0.79 [0.59, 1.04][b]	NA	0.71 [0.49, 1.02][b]
Hist.TB (Historical persistence -more than 1 year positive- of TB on farm in previous years (2003-2006))	4.23 [2.68, 6.69][a]	NA	3.84 [2.32, 6.36][a]
σ_j^2 (SE)	0.86 (0.93)	0.44 (0.66)	0.87 (0.93)
AIC	740.2	712.2	512.4
AUC	0.824	0.767	0.852

[a] significant using 95% CI; [b] significant using 90% CI; NA = not applicable; SE = standard error; AIC = Akaike information criterion; AUC = Area under the curve of the receiver operating characteristic (ROC) curve.

DISCUSSION

This study is relevant because, to the best of our knowledge, no detailed risk factor analysis (at farm and game management unit level) for TB occurrence, incidence and persistence at the wildlife-livestock interface, and on such a large spatial scale, has previously been conducted. This is also the first epidemiological evidence relating cattle TB outbreaks in Spain with wildlife abundance and proximity.

The most interesting result is the significant association between factors related with the proximity, management and abundance of wild ungulates and the TB status of cattle farms. In this regard, the most significant risk factor for TB occurrence, new infection or persistence was having two or more fenced hunting estates within a 2-km radius around the farm. Conversely, a high number of open hunting estates was found to be a protective factor for TB persistence. It is important to note that fenced hunting estates usually have much higher wild ungulate densities than open ones (Acevedo et al., 2008), and this high abundance certainly favours higher contact rates among TB reservoirs, increasing TB transmission (Vicente et al., 2007). Also, many open estates are primarily devoted to small game, which does not share TB susceptibility with cattle. Moreover, there was an inverse correlation (Spearman Rho = -0.25) between the number of fenced hunting estates and the number of open hunting estates around a farm, evidencing that, overall, land use is clearly segregated towards fenced big game (western CR) or open small game (eastern CR), conveying different risks for cattle. This East-to-West pattern in TB risk for cattle was previously evidenced at municipality level by Rodriguez-Prieto et al. (2012). This may explain, at least in part, why fenced hunting states appear to significantly increase the risk of TB on nearby cattle farms whereas open hunting estates may be protective in some situations (i.e. for TB persistence). Similarly, there was a significant association between the risk of new TB infection and (i) a high number of hunted deer per hunting estate within a 2-km radius around the farm and (ii) a high (>497 meters) minimum distance to open big game hunting estates, underlining the importance that wild ungulates abundance and proximity to cattle farms have in TB introduction into TB-free farms. Again, note that there was an slightly inverse correlation (Spearman Rho = -0.02) between the minimum distance to open hunting states and the minimum distance to fenced hunting estates, which implies that, if a cattle farm is far away from an open hunting estate, it may be near to a fenced hunting state and vice versa. There was also a strong positive correlation (Spearman Rho = 0.734) between the number of hunted deer and the number of hunted wild boar per hunting estate, which suggests that, even if the model reveals that a high number of hunted deer, which is a proxy for deer abundance, is a risk factor for new TB infection, this association may be similar with wild boars as suggested in previous studies (Aranaz et al., 1996; Gortázar et al., 2005). Both species have been suggested as true TB reservoirs in South-Central Spain (Vicente et al., 2006; Gortázar et al., 2012). The models also indicate that extensive production systems (i.e. beef and bullfighting cattle farms), in which contact between cattle and wild reservoirs is more likely, are at 3.03, 3.01 and 3.69 times higher risk of TB occurrence, new infection and persistence, respectively, than dairy farms. As all those factors are indicators related to wild ungulates abundance, management and potential contact with cattle, we may conclude that wild ungulates reservoirs are playing a crucial role in TB risk in CR. These results are in agreement with previous studies in other countries and territories and reinforce the need to consider wildlife reservoirs in TB eradication programs (Nishi et al., 2006; O'Brien et al., 2006; Gortázar et al., 2011).

Another interesting result was that an increase in farm size (i.e. number of cattle on the farm) resulted in a 1.99, 1.33 and 2.19 higher risk of TB occurrence, new infection and persistence, respectively. This result is in agreement with previous studies and may be related either to a greater probability of infection in large cattle herds or an increased probability of having a false positive reactor (Humblet et al., 2009). In fact, large herds generally have a higher probability of sharing pastures with wild ungulates or to have contact with adjacent herds which may increase the risk of TB transmission as suggested by Humblet et al. (2010).

Models also reveal a significant but negative association between the average number of cattle leaving the farm per year and TB occurrence and persistence. This may be explained, at least in part, either by (i) the restricted/lower number of outgoing movements in TB-positive farms, (ii) more intensive sanitary control measures imposed on farms with an intensive trade, or (iii) the continuous replacement of cattle (slaughter of old individuals which are more likely to be TB-infected) in those farms with high number of outgoing movements per year. Conversely, a high average number of outgoing cattle shipments per year seems to be a risk factor for new TB infection. This result may be due to higher surveillance pressure in farms with a high number of movements; however further studies should be conducted to more precisely address the specific role that cattle movements may have in TB infection on cattle farms in CR, particularly because previous studies have identified movement of cattle as a significant risk factor for TB infection (Gopal et al., 2006; Humblet et al., 2010).

The significance of the historical persistence (more than 1 year positive) of TB on farm in previous years (2003-2006) for TB occurrence and persistence may indicate that (i) the same risk factors, especially wildlife, persist or even are becoming more risky with time (e.g. increasing hunting industry) in the areas where these farms are located; (ii) farms with previous presence of TB-positive animals may remain infected despite testing and slaughtering of positive individuals due to the low sensitivity of the test (i.e. some results may be false negatives); and (iii) continuous IDTT testing on the farm has increased the number of false positives (i.e. reactors). In any case, the historical persistence of TB on the farm during previous years (2003-2006) seems to increase the risk of TB occurrence and TB persistence on the farm by around 4 times. This result may indicate that, despite the slaughter of positive animals and the rigorous surveillance and control measures implemented on the farm by the eradication programme, a positive farm is likely to become positive again in the near future (1-8 years). Although this finding may imply a long-term effectiveness of the current eradication program in Spain, it is also suggesting the need to look for alternative, short-term and more cost-effective strategies. A proposed alternative strategy may be “individually designed” biosecurity programmes to mitigate risks from wildlife or the implementation of risk-based control and eradication strategies. Results of our study reveal that interventions should prioritise/target high risk areas, which are mostly concentrated in the south-west part of Ciudad Real. Moreover, the large value of the variance of the random effects (σ_j^2) in the three models (Table 1) indicates strong disease clustering within municipalities (i.e. farms within a municipality are very similar) and large variation between municipalities (i.e. farms are very different between municipalities), which suggests that TB interventions would be much more cost-effective if primarily focused in high risk municipalities.

Finally, the models reveal that cattle farms near to goat farms are at a higher risk of TB infection whereas cattle farms near to sheep farms are at a lower risk of TB infection. These results suggest, on the one hand, an important role of goats as TB reservoirs and, on the other, certain resistance of sheep to TB infection, in agreement with suggestions from previous studies (Aranaz et al., 1996; Marianelli et al., 2010; Gortázar et al., 2011).These results clearly highlight the need to compulsorily incorporate goat farms in sanitary campaigns in order to reduce TB on cattle farms and to improve the overall effectiveness of the TB eradication programme.

The statistical models used here (i.e. multilevel logistic regression) and the detailed/high quality data used were considered the most appropriate to address the scientific questions and evaluate the associations between the studied factors and TB occurrence, new infection and persistence in CR. Nevertheless, there is a lot of work still to do to completely understand the complexity of TB transmission and the factors that contribute to hampering TB eradication programmes worldwide. There is a need not only to better understand the interaction between wildlife reservoirs and domestic animals but also to investigate the most cost-effective ways to prevent infectious contacts between them. We also need to further understand how changes in management of farms or hunting estates may affect those wild ungulates-livestock interactions.

It is important to highlight again that an increase in wild ungulates/cattle density and proximity in a TB-endemic area may increase the number of contacts between susceptible and infected animals, which ultimately may lead to an increase in TB transmission. Results presented here have reinforced this idea. This aspect has to be considered particularly in those hunting estates that are intentionally increasing the density of wild ungulates for hunting purposes, but without the implementation of adequate biosecurity measures to prevent and control TB transmission. The proper management of the hunting areas and national parks seems crucial to prevent further TB transmission and, ultimately, to better control and eradicate TB in domestic and susceptible wildlife populations. These results are of value for supporting policies in Central and Southern Iberia, and probably in other areas where epidemiological factors such as Mediterranean conditions, extensive cattle breeding systems and increasing big game activity and industry concur, and also for other domestic species such as goats and free-roaming pigs, for which more research is needed (Di Marco et al., 2012).

ACKNOWLEDGEMENTS

This study has been funded by research projects FAU2008-0004-C03 and JCCM PPEII 10-0262-7673. This is a contribution to EU project number 278976 (ANTIGONE) Beatriz Martínez-López holds a Juan de la Cierva post-doctoral contract (JCI-2011-10724) and Victor Rodríguez-Prieto holds a FPU pre-doctoral scholarship, both funded by the Spanish Ministry of Economy and Competitiveness. Authors would like to acknowledge everyone who contributed to the wildlife and hunting data collection over the years as well as provision of data for cattle farms and hunting estates by the Regional Veterinary and Forestry Authorities of Castile-La Mancha [Consejería de Agricultura de Castilla-La Mancha] and the Ministry of Agriculture, Food and Environment.

REFERENCES

Acevedo, P., Ruiz-Fons, F., Vicente, J., Reyes-García, A.R., Alzaga, V. and Gortázar, C. (2008). Estimating red deer abundance in a wide range of management situations in Mediterranean habitats. J. Zool. 276, 37–47

Alexander, K.A., Pleydell, E., Williams, M.C., Lane, E.P., Nyange, J.F. and Michel, A.L. (2002). Mycobacterium tuberculosis: an emerging disease of free-ranging wildlife. Emerg. Infect. Dis. 8, 598-601

Allepuz, A., Casal, J., Napp, S., Saez, M., Alba, A., Vilar, M., Domingo, M., Gonzalez, M.A., Duran-Ferrer, M., Vicente, J., Alvarez, J., Muñoz, M. and Saez, J.L. (2011). Analysis of the spatial variation of bovine tuberculosis disease risk in Spain (2006–2009). Prev. Vet. Med. 100, 44–52

Aranaz, A., Liébana, E., Mateos, A., Domínguez, L., Vidal, D., Domingo, M., Gonzólez, O., Rodríguez-Ferri, E.F., Bunschoten, A.E., van Embden, J.D. and Cousins, D. (1996). Spacer oligonucleotide typing of Mycobacterium bovis strains from cattle and other animals: a tool for studying epidemiology of tuberculosis. J. Clin. Microbiol. 34, 2734-2740

Belsley, D. A., Kuh, E., and Welsch, R. E. (1980). Regression Diagnostics; Identifying Influential Data and Sources of Collinearity, New York: John Wiley & Sons

Di Marco, V., Mazzone, P., Capucchio, M.T., Boniotti, M.B., Aronica, V., Russo, M., Fiasconaro, M., Cifani, N., Corneli, S., Biasibetti, E., Biagetti, M., Pacciarini, M.L., Cagiola, M., Pasquali, P. and Marianelli, C. (2012). Epidemiological significance of the domestic black pig (Sus scrofa) in maintenance of bovine tuberculosis in Sicily. J. Clin. Microbiol. 50, 1209-1218

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

Gopal, R., Goodchild, A., Hewinson, G., De la Rua-Domenech, R. and Clifton-Hadley, R. (2006). Introduction of bovine tuberculosis to north-east England by bought-in cattle. Vet. Rec. 159, 265–271

Gortázar, C., Delahay, R. J., McDonald, R. A., Boadella, M., Wilson, G. J., Gavier-Widén, D. and Acevedo, P. (2012). The status of tuberculosis in European wild mammals. Mammal Review 42, 193-206

Gortázar, C., Vicente, J., Boadella, M., Ballesteros, C., Galindo, R.C., Garrido, J., Aranaz, A. and de la Fuente, J. (2011). Progress in the control of bovine tuberculosis in Spanish wildlife. Vet. Microbiol. 151, 170-178

Gortázar, C., Vicente, J., Samper, S., Garrido, J.M., Fernandez-de-Mera, I.G., Gavin, P., Juste, R.A., Martin, C., Acevedo, P., De La Puente, M. and Höfle, U. (2005). Molecular characterization of Mycobacterium tuberculosis complex isolates from wild ungulates in south-central Spain. Vet. Res. 36, 43-52

Humblet, M.-F., Boschiroli, M.L. and Saegerman, C. (2009). Classification of worldwide bovine tuberculosis risk factors in cattle: a stratified approach. Vet. Res. 40, 50

Humblet, M.-F.,Gilbert, M., Govaerts, M., Fauville-Dufaux, M., Walravens, K. and Saegerman, C. (2010). New Assessment of Bovine Tuberculosis Risk Factors in Belgium Based on Nationwide Molecular Epidemiology. J. Clin. Microbiol. 48, 2802-2808

MAGRAMA (2012). Technical and financial final report of the National programme for bovine tuberculosis in Spain during 2011. In Spanish. Available at: http://rasve.mapa.es/Recursos/Ficheros/Historico/62_RESUMEN%20ANUAL%20TB%202011.pdf

Marianelli, C., Cifani, N., Capucchio, M.T., Fiasconaro, M., Russo, M., La, M.F., Pasquali, P. and Di Marco, V. (2010). A case of generalized bovine tuberculosis in a sheep. J. Vet. Diagn. Invest 22, 445-448

Nishi, J.S., Shury, T. and Elkin, B.T. (2006). Wildlife reservoirs for bovine tuberculosis (Mycobacterium bovis) in Canada: strategies for management and research. Vet Microbiol. 112, 325-338

O'Brien, D.J., Schmitt, S.M., Fitzgerald, S.D., Berry, D.E. and Hickling, G.J. (2006). Managing the wildlife reservoir of Mycobacterium bovis: the Michigan, USA, experience. Vet. Microbiol. 112, 313-323

Phillips, C.J.C., Foster, C.R.W., Morris, P.A. and Teverson, R., (2003). The transmission of Mycobacterium bovis infection to cattle. Res. Vet. Sci. 74, 1–15

Porphyre, T., Stevenson, M. A. and McKenzie, J. (2008). Risk factors for bovine tuberculosis in New Zealand cattle farms and their relationship with possum control strategies. Prev. Vet. Med. 86, 93–106

R Development Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/

Reviriego Gordejo, F.J. and Vermeersch, J.P. (2006). Towards eradication of bovine tuberculosis in the European Union. Vet. Microb. 112, 101-109

Rodríguez-Prieto, V., Martínez-López, B., Barasona, J.A., Acevedo, P., Romero, B., Rodriguez-Campos, S., Gortázar, C., Sánchez-Vizcaíno, J.M. and Vicente, J. (2012). A Bayesian approach to study the risk variables for tuberculosis occurrence in domestic and wild ungulates in South Central Spain. BMC Vet. Res. 8:148

Schiller, I., Waters, W.R., Vordermeier, H.M., Jemmi, T., Welsh, M., Keck, N., Whelan, A., Gormley, E., Boschiroli, M.L., Moyen, J.L., Vela, C., Cagiola, M., Buddle, B.M., Palmer, M., Thacker, T. and Oesch, B. (2011). Bovine tuberculosis in Europe from the perspective of an officially tuberculosis free country: trade, surveillance and diagnostics. Vet. Microbiol. 151, 153-159

Smith, N.H., Berg, S., Dale, J., Allen, A., Rodriguez, S., Romero, B., Matos, F., Ghebremichael, S., Karoui, C., Donati, C., Machado Ada, C., Mucavele, C., Kazwala, R.R., Hilty, M., Cadmus, S., Ngandolo, B.N., Habtamu, M., Oloya, J., Muller, A., Milian-Suazo, F., Andrievskaia, O., Projahn, M., Barandiarán, S., Macías, A., Müller, B., Zanini, M.S., Ikuta, C.Y., Rodriguez, C.A., Pinheiro, S.R., Figueroa, A., Cho, S.N., Mosavari, N., Chuang, P.C., Jou, R., Zinsstag, J., van Soolingen, D., Costello, E., Aseffa, A., Proaño-Perez, F., Portaels, F., Rigouts, L., Cataldi, A.A., Collins, D.M., Boschiroli, M.L., Hewinson, R.G., Ferreira Neto, J.S., Surujballi, O., Tadyon, K., Botelho, A., Zárraga, A.M., Buller, N., Skuce, R., Michel, A., Aranaz, A., Gordon, S.V., Jeon, B.Y., Källenius, G., Niemann, S., Boniotti, M.B., van Helden, P.D., Harris, B., Zumárraga, M.J. and Kremer, K. (2011). European 1: A globally important clonal complex of Mycobacterium bovis. Infect. Genet. Evol. 11, 1340-1351

Venables, W.N. and Ripley, B.D. (2002). Modern Applied Statistics with S. New York: Springer (4th ed)

Vicente, J., Höfle, U., Garrido, J.M., Fernández-de-Mera, I.G., Acevedo, P., Juste, R., Barral, M. and Gortázar, C. (2007). Risk factors associated with the prevalence of tuberculosis-like lesions in fenced wild boar and red deer in south central Spain. Vet. Res. 38, 451-464

Vicente, J., Höfle, U., Garrido, J.M., Fernández-de-Mera, Juste, R., Barral, M. and Gortázar, C. (2006). Wild boar and red deer display high prevalences of tuberculosis-like lesions in Spain. Vet. Res. 37, 107-119

VETERINARY PUBLIC HEALTH

MAPPING THE DISTRIBUTION OF *ECHINOCOCCUS GRANULOSUS* IN WALES AND GREAT BRITAIN

H. TEMPLE[*], P.H. JONES AND A. BROUWER

SUMMARY

Echinococcus granulosus is a cestode zoonotic parasite whose life cycle involves dogs and other canids as the definitive hosts, and sheep and other ruminants as the intermediate hosts. Humans may also act as 'accidental' intermediate hosts. In this study, cattle have been used as a sentinel host to investigate the distribution of *E. granulosus* in Wales, with relevance for Great Britain. All UK registered cattle slaughtered at selected abattoirs were included in the analysis. 757 of 893,379 cattle (0.085%) were identified as having hydatid cysts, consisting of 91 males and 666 females. The movement records of all the cattle slaughtered were obtained and used to identify the most likely locations of infection. The investigation demonstrated that *E. granulosus* is not confined to the South Powys area of Wales, as commonly suggested, but is widespread, appearing to be present throughout Wales and England, with at least some locations in Scotland.

INTRODUCTION

Hydatid disease (hydatidosis, cystic echinococcosis) is a parasitic infection caused by the larval stages of the cestode *Echinococcus granulosus.* It is an important infection of livestock worldwide, and one of the most widespread parasitic zoonoses (Lloyd et al., 1991; Holcman & Heath, 1996; Rinaldi et al., 2008; Andresiuk et al., 2009; Ibrahim, 2010; Guo et al., 2011).

Dogs and foxes, the normal definitive hosts (Walters, 1978; Lloyd, et al. 1991; Andresiuk et al., 2009), are infected through the ingestion of offal from infected animals containing hydatid cysts, the larval stage of *E. granulosus* (Rinaldi et al., 2008; Andresiuk et al., 2009). The tapeworm grows in the intestine of the definitive hosts and eggs are released from segments passed in the faeces (Walters, 1978; Andresiuk et al., 2009). Sheep are the main intermediate hosts, although cattle, horses and pigs can become infected via ingestion of eggs on the pasture (Walters, 1978; Lloyd et al., 1991; Romig et al., 2006; Rinaldi et al., 2008; Andresiuk et al., 2009). Humans can become 'accidental' intermediate hosts following inadvertent ingestion of eggs (Andresiuk et al., 2009; Mastin et al., 2011). The eggs hatch in the intestine, freeing oncospheres which go on to penetrate the intestinal wall and are carried

---

[*] Hannah Temple, University of Liverpool, Leahurst Campus, Neston, Cheshire CH64 7TE
Email: H.Temple@student.liverpool.ac.uk or Arjen.Brouwer@wales.gsi.gov.uk

via the bloodstream to various parts of the body (Walters, 1978; Andresiuk et al., 2009; Brouwer & Willson, 2009). The main locations for cyst development are the liver and lungs, and rarely the spleen, heart and even the brain, where the parasite grows as a fluid filled cyst, known as a hydatid cyst, the metacestode stage (Breijo et al., 2008; Andresiuk et al., 2009; Brouwer & Willson, 2009). The cysts can reach up to several centimetres in diameter. The resulting clinical disease in humans, human hydatid disease, causes considerable morbidity and mortality worldwide (Breijo et al., 2008; Andresiuk et al., 2009; Mastin et al., 2011). In Great Britain (GB), the predominant strain is the sheep strain (G1). The parasite predominantly follows a sheep/dog lifecycle, and cysts in cattle tend to be infertile (Thompson, 1995; Romig et al., 2006; Andresiuk et al., 2009; DEFRA, 2011). There is also a horse adapted strain (G4) present in GB, involving a dog/horse lifecycle (Andresiuk et al., 2009; DEFRA, 2011).

The patterns of human and animal infection with hydatids vary considerably between countries worldwide (Walters, 1978). With regard to human hydatid disease in GB, it occurs in many counties and the precise proportion contracted outside of GB is unknown. Historically, human cases have been associated with the farming areas of South Powys, Herefordshire and some of the Scottish Islands. These areas had large sheep populations often worked by sheep dogs, which had free access to carcases or were fed raw meat and/or offal from fallen stock or on-farm slaughter (Romig et al., 2006; Brouwer & Willson, 2009).

Within Wales, the incidence and distribution of human hydatid disease has changed a lot over time. Prior to the 1960s, reports of human hydatid disease were primarily associated with the industrial valleys of Gwent and Glamorgan in South Wales, whereas between the 1960s and 1990s reports came from Powys, mid-Wales (particularly Brecknockshire region, South Powys) (Brouwer & Willson, 2009; Mastin et al., 2011). From 1974 to 1983 the incidence of new cases in Wales was 0.27 per 100,000, with the highest incidence in Powys, especially Brecknockshire (7 per 100,000) (Lloyd et al., 1991; Brouwer & Willson, 2009). A hydatid disease eradication scheme was run in South Powys from 1983 to 1989. The scheme was based on voluntary, supervised dosing of dogs with praziquantel and education of farmers. After the termination of the dog dosing scheme in 1989, a health education campaign was initiated over a larger area, targeting school children and farmers (Brouwer & Willson, 2009; Mastin et al., 2011).

In 2005, the prevalence of *E. granulosus* in dogs raised concerns that accidental infection of humans may emerge once again, especially in younger age groups. An all-Wales awareness campaign was implemented by the Welsh Government, with information packs sent to all veterinary clinics, doctors, dentists and pharmacists in Wales (Brouwer & Willson, 2009).

The incidence of autochthonous hospital-diagnosed echinococcosis in humans in recent years has declined to negligible levels (Mastin et al., 2011). However, from 2000 to 2010, there have been 142 cases identified in England and Wales (Health Protection Agency, 2009). In 2010 there were 6 cases in England and Wales and 1 case in Scotland. Of the 7 cases, 5 were thought to have contracted the disease abroad (DEFRA, 2011). Without continued vigilance and public awareness of the disease, there is a risk that the number of human cases will rise again.

The main aim of this study is to map the geographical distribution of *E. granulosus*, using cattle as sentinel hosts and indicators of risk for possible human infection, as part of the

Welsh Government's national disease awareness efforts. Cattle have been investigated instead of sheep, as sheep are slaughtered much younger, often before the hydatid cysts have developed to a recognisable size. Additionally, the gross pathology in cattle is less likely to be mistaken for differential diagnoses (Brouwer & Willson, 2009).

MATERIALS AND METHODS

Data collection

Data were collected by the Meat Hygiene Service (FSA Ops) on behalf of the Welsh Government between 1st October 2010 and 30th September 2011 at 19 selected abattoirs. The abattoirs were selected based on the percentage of the Welsh cattle population, defined as cattle born in Wales and still on Welsh premises at 12 and 18 months of age, slaughtered at these premises. During the study period, the selected abattoirs slaughtered 78% of the specified Welsh cattle population going for slaughter in GB. These abattoirs also take cattle from all over the UK. These animals were also included in this study; therefore potential hotspots of *E. granulosus* were investigated outside Wales, particularly England.

Cyst identification was carried out by Meat Hygiene Service (FSA Ops) meat inspectors during their routine inspections, following specific refresher training on hydatidosis. Following identification of a hydatid cyst, the official identity of the animal was recorded, as well as information regarding location, consistency and number of cysts.

Data management

The movements of all cattle slaughtered at these premises were obtained from the Animal Health Veterinary Laboratories Agency (AHVLA). In addition, the date of birth, breed and sex of all of the cattle slaughtered were obtained.

Data analysis

Analysis has been carried out on the dataset looking at movements of hydatid identified cattle and the prevalence of hydatidosis per postcode areas. The raw data were stored in a MySQL database (Oracle Corporation, Redwood Shores, CA) and imported into Stata 12 (StataCorp., College Station, TX) for analysis. For the purpose of this study, non-UK registered cattle at birth were not included as it was not possible to determine if they were infected prior to entering the UK. Six of 3,577 non-UK registered cattle were identified with hydatid cysts.

The movement data were analysed in two ways; firstly by mapping the geographic distribution of cattle identified with hydatid cysts, by location, and secondly by calculating and mapping the prevalence by postcode area. Fifty-seven cattle had some data missing from their movement records, leaving 700 with complete histories. Many cattle made short movements consisting of zero or one days, often to markets or auctions, and sometimes returning to the same premises. These movements were considered to be unlikely sources of infection and discounted when assessing the locations where disease might have been contracted. Any locations hydatid identified cattle were translocated to within 170 days (6 months) prior to slaughter were discounted from the study due to the time required for any cysts to develop to a visibly recognisable size. There is very little information regarding the growth rate of cysts in the literature. It can be slow and variable, and appears dependent on a number of factors including the strain of parasite, species and breed of host (both definitive and intermediate) and the degree of infection. Depending on these factors, cysts can increase in diameter by between 1 and 5 cm per year (Thompson, 1995). In sheep, 6-9 months post infection the cysts measured 2-5 mm in diameter, and at 12 months the cysts were more easily identifiable (Heath, 1980). Six months was decided as a reasonable minimum time period required for the cysts to develop in cattle.

To map out the distribution of hydatid disease, relevant cattle were categorised depending on the number of movements carried out. Maps showing the movements of the cattle were produced in Quantum GIS (QGIS) version 1.7.4 (Open Source Geospatial Foundation, Beaverton, OR)

For determining the prevalence of hydatid disease, only cattle that had never left a postcode area were included in this analysis. Queries were run in MySQL to determine which cattle had remained within a postcode area and the resulting data were imported into Stata to calculate prevalence.

RESULTS

Hydatid cysts

During the 12-month study period 757 (0.085%) cattle were identified as having hydatid cysts out of 893,379 cattle slaughtered at the 19 selected abattoirs.

Movement analysis

The movements of hydatid cattle have been split into the number of movements undertaken and mapped out in Fig. 1. Figure 1 shows cattle that spent their entire life on one premises in solid black dots. These premises are regarded as the most likely source of infection. The spot size for the locations is relative to the stay length at the premises. There were 373 cattle that had been on only 1 holding; 53% of all cattle with complete movement histories. The cattle that had been on 2 locations (n=202) are marked in dark grey as per the legend. It can be hypothesised that the location at which the cattle dwelt the longest is most likely to be the source of infection, given equal risk of becoming infected, particularly when the duration at the second location was very short. Figure 1 contains 88 cattle that had been on 3 premises during their life, and 37 cattle that had been on 4 or more premises prior to slaughter.

The map shows that *E. granulosus* is not confined to the South Powys and Herefordshire areas, with which it is historically associated. The distribution is much wider and the infection appears to be present throughout Wales and England, with evidence of some locations in Scotland. The results are focused in and around Wales, and it must be remembered that the abattoirs were selected based on the percentage of Welsh cattle slaughtered, so there is a recognised and intended bias towards Wales.

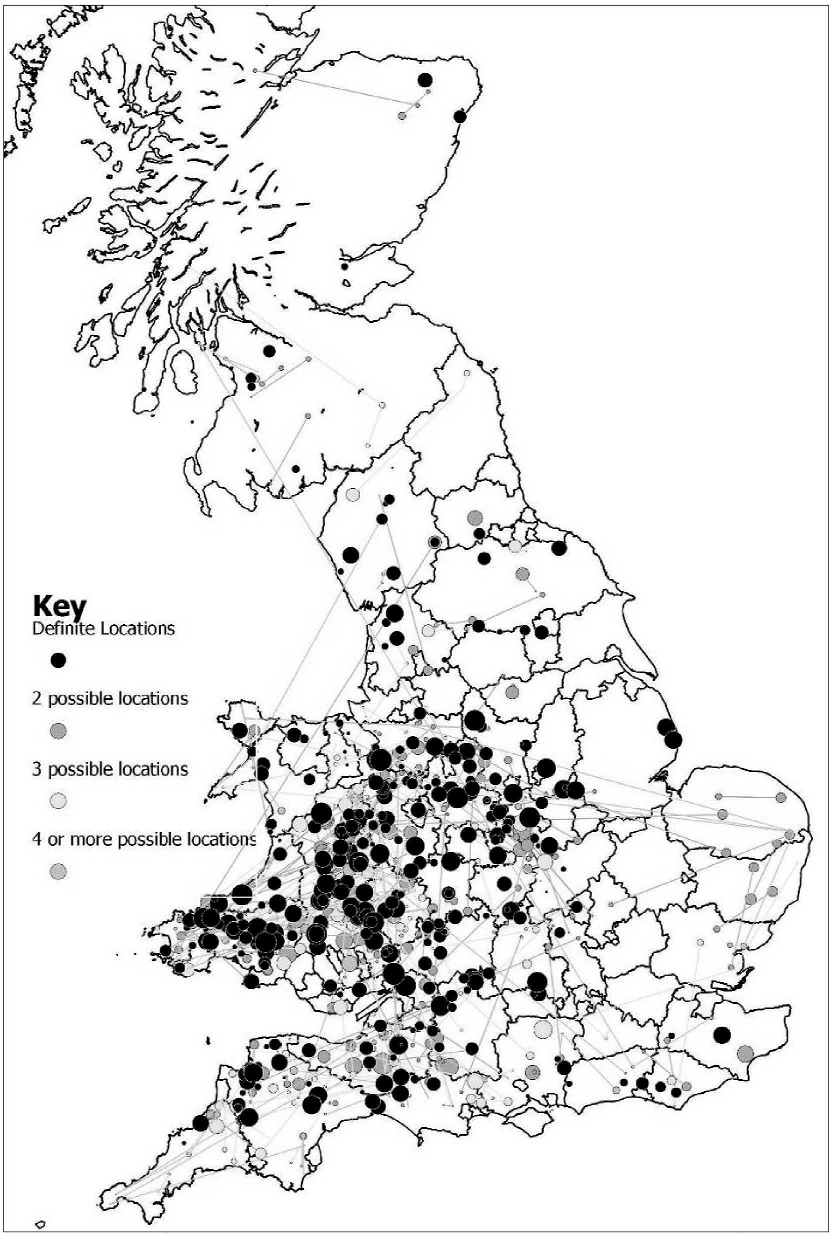


Fig. 1 Map showing the locations and movements of hydatid identified cattle excluding translocations of 0 and 1 days, and locations 170 days prior to slaughter. The spot size is relative to the length of time at the location

Prevalence of hydatidosis by postcode areas

Only cattle that were slaughtered at the 19 selected abattoirs and did not leave a postcode area during the relevant time frame were included to calculate the prevalence of hydatidosis by postcode area. The prevalence has been mapped out and coloured according the number of hydatid cattle per 100,000 cows: the darker the colour, the higher the prevalence (Fig. 2).

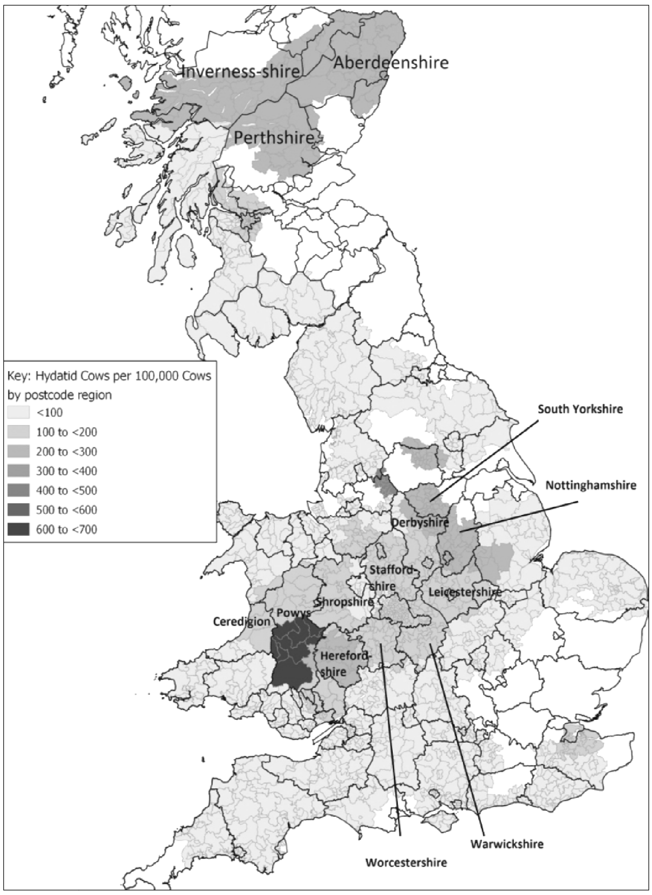


Fig. 2 Estimated prevalence of hydatidosis per 100,000 cattle slaughtered by postcode area

The map corroborates the findings of the map in Fig. 1, with *E. granulosus* present throughout Wales and England. There is a higher prevalence in and around mid-Wales and the Midlands of England, as well as in Aberdeenshire and Perthshire in Scotland. A small hotspot appears to the East of Manchester in the Oldham postcode with a prevalence of 481 per 100,000 cattle but there are relatively large 95% confidence intervals (0, 1423.06), which is most likely due to few slaughtered animals being from there.

DISCUSSION

Human hydatid disease is an important public health concern, and one of the objectives of this study was to identify areas of possible increased risk to humans in Wales and Great Britain (Brouwer & Willson, 2009; Ibrahim, 2010). Due to the higher human incidence in mid Wales, the South Powys Hydatid Control Programme was implemented, providing free preventative deworming treatment to farmers to treat their dogs. The programme was deemed a success with a marked reduction in human and ovine cases within the treatment area. Data for the period 1984 to 1990 showed the incidence of new human cases fell from 0.27 to 0.22 per 100,000 in Wales, and specifically in Powys the numbers fell from 3.89 to 2.28 per 100,000 (Brouwer & Willson, 2009). The number of human cases in Wales, especially Powys, has been declining since the implementation of various control measures over the years. However, without the implementation of these control programmes the numbers of human cases might have continued to rise or remain unacceptably high.

Continued monitoring of the disease in humans and livestock, as well as public awareness of the disease, is advisable to ensure the number of human cases does not rise again. Through tracking the movements of hydatid positive cattle, areas of potentially higher risk for human transmission can be identified, allowing control/prevention programmes to be targeted. The results of this investigation have shown *E. granulosus* to be very widely distributed and not confined to South Powys and Herefordshire as historical data from human infection may suggest (Brouwer & Willson, 2009). This study's results are not representative for England or Scotland and nationally relevant studies should be carried out in these countries to show the geographical distribution throughout GB; this study has highlighted potential hotspots of disease outside of Wales, which warrant investigation. These results highlight the need for continued vigilance and routine anthelminthic treatment of dogs with appropriate products in Wales and other parts of GB, to prevent transmission to humans.

Investigating sheep for this study was considered but as they are slaughtered much younger, and often before the hydatid cysts have developed to a visible size, cattle were chosen. Also, compared to sheep the gross pathology in cattle is less likely to be mistaken for differential diagnoses (Heath, 1980; Thompson, 1995; Brouwer & Willson, 2009). Cattle have provided a good alternative as a sentinel host although the population densities and distribution of cattle and sheep in GB must be born in mind when interpreting results.

This investigation relied on cooperation from meat hygiene inspectors at the 19 selected abattoirs. The identification of the cysts was not carried out by the same person; hence measurement bias may have occurred. Diagnosis was carried out only by visual identification, incision and palpation as per routine meat inspection procedures, and this is likely to have led to an underestimation of cases. All meat inspectors carried out specific refresher training, although there was no measurement to determine how effective the training was.

ACKNOWLEDGEMENTS

The authors would like to thank Richard Hughes at the Welsh Government for facilitating the data and providing help in preparing data for further analysis, and Jon Weston and Chris Cotterill at AHVLA for their help in providing the cattle movement data. The Welsh Government provided funding to enable the training of and recording by FSA staff. We gratefully acknowledge funding from the Wellcome Trust for this MSc project.

REFERENCES

Andresiuk, M.V., Gordo, F.P., Bandera, C.C., Elissondo, M.C., Dopchiz, M. and Denegri, G. (2009). Echinococcus granulosus: biological comparison of cattle isolates from endemic regions of Argentina and Spain. Rev. Argent. Microbiol. 41, 218-225

Breijo, M., Anesetti, G., Martínez, L., Sim, R.B. and Ferreira, A.M. (2008). Echinococcus granulosus: the establishment of the metacestode is associated with control of complement-mediated early inflammation. Exp. Parasitol. 118, 188-196

Brouwer, A. and Willson, C. (2009). The history of Hydatid control in Wales. Government Vet. J. 20, 32-35

DEFRA (2011). Zoonoses Reports UK 2010 and 2011

Guo, Z.H., Kubo, M., Kudo, M., Nibe, K., Horii, Y. and Nonaka, N. (2011). Growth and genotypes of *Echinococcus granulosus* found in cattle imported from Australia and fattened in Japan. Parasitol. Int. 60, 498-502

Health Protection Agency (2009). Hydatid Disease (Echinococcosis) Epidemiology Data Available: http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/HydatidDisease/EpidemiologicalData/ [2012, August]

Heath, D.D. (1980). The development of Echinococcus granulosus cysts in sheep. N. Z. Zool. 7, 605-606

Holcman, B. and Heath, D.D. (1996). The early stages of Echinococcus granulosus development. Acta. Trop. 64, 5-17

Ibrahim, M.M. (2010). Study of cystic echinococcosis in slaughtered animals in Al Baha region, Saudi Arabia: Interaction between some biotic and abiotic factors. Acta. Trop. 113, 26-33

Lloyd, S., Martin, S.C., Walters, T.M.H. and Soulsby, E.J.L. (1991). Use of sentinel lambs for early monitoring of the South Powys [Wales, UK] Hydatidosis Control Scheme: prevalence of Echinococcus granulosus and some other helminths. Vet. Rec. 129, 73-76

Mastin, A., Brouwer, A., Fox, M., Craig, P., Guitián, J., Li, W. and Stevens, K. (2011). Spatial and temporal investigation of Echinococcus granulosus coproantigen prevalence in farm dogs in South Powys, Wales. Vet. Parasitol. 178, 100-107

Rinaldi, L., Maurelli, M.P., Veneziano, V., Capuano, F., Perugini, A.G. and Cringoli, S. (2008). The role of cattle in the epidemiology of Echinococcus granulosus in an endemic area of southern Italy. Parasitol. Res. 103, 175-179

Romig, T., Dinkel, A. and Mackenstedt, U. (2006). The present situation of echinococcosis in Europe. Parasitol. Int. 55, S187-S191

Thompson, R.C.A. (1995). Echinococcus and Hydatid Disease. 1st edn. CAB International, Wallingford

Walters, T.M.H. (1978). Hydatid disease in Wales. 1. Epidemiology. Vet. Rec. 102, 257-259

COXIELLA BURNETII IN SMALL RUMINANTS: ASSESSING THE PUBLIC HEALTH RISK

T.C.M. DEWÉ*, A. NIGSCH AND K.D.C. STÄRK

SUMMARY

Coxiella burnetii causes Q fever, a zoonosis that can have serious consequences. Most human cases are associated with small ruminant parturition, as was the case in the large Netherlands outbreak of 2007-2010. However, a quantitative assessment of the public health risk attributable to *C. burnetii* infection in small ruminants has not previously been performed. This study aimed to do so for Switzerland, using a model developed in Palisade @RISK Version 5.7. Goat flocks were shown to represent a greater risk of *C. burnetii* shedding than sheep flocks, and the number of *C. burnetii* released by an infected sheep or goat flock was estimated to be in the order of 10^{12} organisms. The risk to individuals was highest for farmers and declined over a distance of five kilometres from a shedding farm. The model was most sensitive to the *C. burnetii* concentration in small ruminant placental tissue. Other influential inputs were the within-flock shedding prevalence, flock infection prevalence, and flock size.

INTRODUCTION

Q fever is a zoonosis caused by the bacteria *Coxiella burnetii*, which can infect a broad range of species. Q fever in people ranges from asymptomatic infection or mild flu-like illness to rare serious sequelae such as pneumonia, endocarditis and chronic fatigue (Marrie, 1995). Its common clinical manifestation in domestic ruminants is a reproductive syndrome (Maurin & Raoult, 1999). It is an occupational disease of agricultural and abattoir workers (Garner et al., 1997) but can also present a public health risk through airborne transmission. Human outbreaks have mostly been associated with periparturient sheep and goats (EFSA, 2010), often after dry, windy weather (Hawker et al., 1998; Tissot-Dupont et al., 1999). Recently, a large outbreak in the Netherlands affected at least 3500 people, with an unprecedented impact on public health and the small ruminant farming industry (Roest et al., 2010).

C. burnetii has a complex multi-phase life cycle with both intracellular and extracellular forms (Maurin & Raoult, 1999). The bacteria multiply in the foetal placenta (Roest et al., 2012) and are shed in vast numbers during abortions and normal parturition of small ruminants in particular (Welsh et al., 1958). *C. burnetii* has been identified in the placenta, foetal fluids, and aborted and healthy foetuses of sheep and goats (Abinanti et al., 1953;

---

* Tamsin Dewé, Royal Veterinary College, Royal College Street, London NW1 0TU, UK. Email: tamsindewe@gmail.com

Welsh et al., 1958; Roest et al., 2012). *C. burnetii* DNA has also been detected in vaginal mucus, milk and faeces of these animals (Arricau-Bouvery et al., 2003).

C. burnetii is considered highly infectious by inhalation, with fewer than ten organisms thought to be sufficient to cause infection (Tigertt et al., 1961). The infective form of the bacteria is environmentally resistant (Maurin & Raoult, 1999) and the organism has been detected in air and dust samples (Welsh et al., 1958). The precise role of environmental conditions including wind speed and relative humidity in facilitating *C. burnetii* spread is yet to be fully explored (van der Hoek et al., 2011), however, some risk factors have been identified. These include living on or within close proximity to an infected farm (Veenstra et al., 2011) and being downwind of parturient flocks (Hawker et al., 1998). Indirect transmission via contaminated bedding and other materials has also been shown (Marrie et al., 1989).

The host immune response to *C. burnetii* involves both humoral and cellular immunity (Maurin & Raoult, 1999). Antibodies are described as long-lasting and seroconversion often occurs, although the timing can be highly variable in pregnant ruminants and may not appear until weeks after parturition (Arricau-Bouvery et al., 2003). Thus, both seropositive and negative animals can be shedding *C. burnetii*. Latent and chronic infections have been documented (Maurin & Raoult, 1999) and infection within a herd may also be clinically silent (Hilbert et al., 2012). The reported seroprevalence of *C. burnetii* in European sheep ranges from 3% to 30% in individuals, and 0% to 74% at herd level; in goats, individual and herd-level seroprevalence is between 7% and 73%, and 0% and 47%, respectively (Guatteo et al., 2011).

The public health impact of the Netherlands outbreak has prompted qualitative assessments of the risk of Q fever elsewhere (ECDC, 2010). Switzerland contains numerous sheep and goat flocks that are kept in close proximity to humans for at least part of the year. Seroprevalence in the general human population is usually between 2% and 10%, although it can be much higher in occupationally exposed groups (EFSA, 2010). In Switzerland, seroprevalence was found to vary between 7% and 32% (Dupuis et al., 1986), although Q fever cases have been reported only rarely. The last large-scale human outbreak in this country resulted in over 400 acute cases and was attributed to movement of sheep flocks past several villages (Dupuis et al., 1987).

To the authors’ knowledge, the human Q fever risk attributable to *C. burnetii*-infected livestock has not been quantitatively assessed. This study aimed to model the overall risk of acute symptomatic Q fever in the average adult living within five kilometres of a sheep or goat flock in Switzerland.

MATERIALS AND METHODS

A quantitative risk assessment (QRA) was performed according to the World Organization for Animal Health (OIE) framework (2004). The hazard was identified as *C. burnetii* shed by Swiss small ruminant flocks and a comprehensive literature review was undertaken (Dewé & Stärk, 2012). Risk pathways were established for release, exposure and consequence assessments, and inputs were parameterised stochastically (Fig. 1) in @RISK Version 5.7 (Palisade, West Drayton, United Kingdom). Critical data needs were identified after a crude sensitivity analysis and addressed by published and unpublished evidence, as outlined below. The combined model was further refined and run over 5000 iterations.

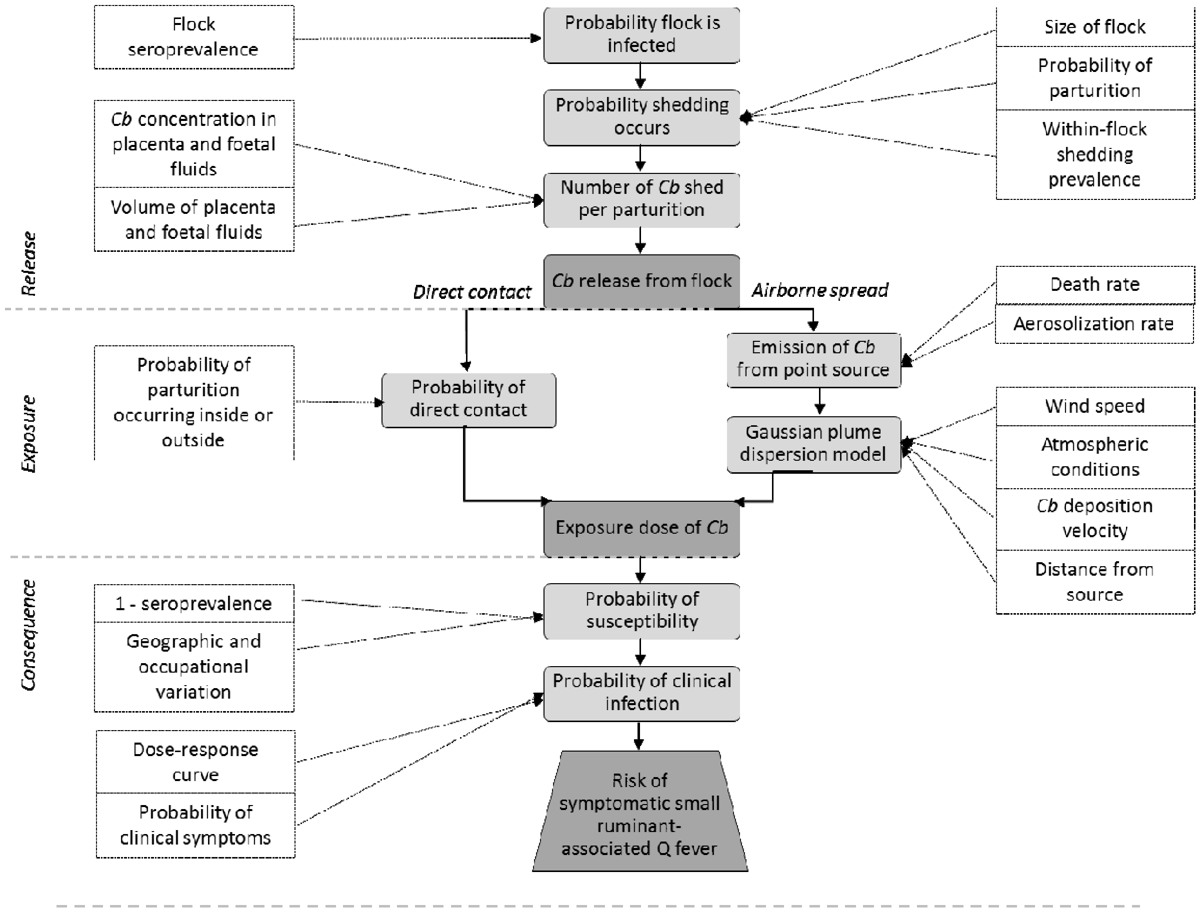


Fig. 1 Risk pathways and main input parameters for the assessment of the risk of small-ruminant associated acute symptomatic Q fever in Switzerland. *Cb: Coxiella burnetii*

Release assessment

The probability of a flock shedding *C. burnetii* and the number of organisms shed during a reproductive season were modelled. Input parameters are described in Table 1; full details, including sources, have been described elsewhere (Dewé & Stärk, 2012).

Table 1. Input parameters for release assessment[a]

	Parameter	Distribution	Sheep value	Goat value
P_i	Flock infection prevalence	Pert alternate parameters (most likely, 2.5%, 97.5%)	Most likely: 0.050, 2.5%: 0.016, 97.5%: 0.113	Most likely: 0.111, 2.5%: 0.049, 97.5%: 0.207
P_p	Probability of parturition	Beta (s +1, n – s +1)	s: 40129, n: 74096	s: 18908, n: 28036
P_s	Within-flock shedding prevalence	Beta alternate parameters (min, max, 50%, 75%)	Min: 0, max: 1, 50%: 0.124, 75%: 0.522	Min: 0, 50%: 0.342, 75%: 0.626, max: 1
N_f	Number of breeding females in a flock	Cumulative distribution	Census data	Census data
ID_{50}	Number of *C. burnetii* in hamster ID_{50}	Uniform (min, max)	Min: 2, max: 4.2	
C_p	*C. burnetii* concentration in placenta ($\log_{10}$/g)	Beta alternate parameters (min, 25%, 50%, 95%)	Log-transformed min and % from combined datasets $1/ID_{50}$ and 2	
C_f	*C. burnetii* concentration in foetal fluids ($\log_{10}$/ml)	Beta alternate parameters (min, 25%, 50%, 95%)	Log-transformed min and % from dataset 3	
V_p	Placental size (g)	Normal (μ, σ)	μ: 579.4, σ: 25.7	μ:539.1, σ: 221.2
V_f	Foetal fluid volume (ml)	Uniform (min, max)	Min: 700, max: 1300	Min: 479, max: 603

[a] Key: %, percentile; min, minimum value; max, maximum value; μ, mean; σ, standard deviation

The probability of shedding (P_f) was determined by:

$$P_f = P_i \cdot \left(1 - \left(1 - P_p \cdot P_s\right)^{N_f}\right) \qquad (1)$$

At flock level, seroprevalence results from an on-going study (Hunninghaus et al., 2012) were used as a proxy for infection. Herd size data were obtained directly from census results (Schwermer, 2012), and truncated at greater than or equal to three animals. Only females of reproductive age (≥ 1 year) were considered. The probability of parturition in a breeding female during a reproductive season was extrapolated from studbook data (Aeschlimann, 2012; Thomann, 2012).

Data on the within-herd seroprevalence and within-herd shedding prevalence were collated from published studies on periparturient animals. Flocks of clinically healthy small ruminants and those experiencing low levels of abortion (< 10%) were considered, as these were considered most representative of endemic conditions in Switzerland. Simple linear regression analysis of the within-herd seroprevalence as the explanatory variable, and within-herd shedding prevalence as the response variable, was performed in R (The R Foundation, Vienna, Austria). No significant association was found for either species (p>0.6, data not shown). The probability of an individual animal shedding at parturition was therefore determined by an empirical distribution of data from these reports.

The total number of *C. burnetii* shed by a flock (N_0) was estimated from the number of shedding animals in a flock (N_s) by:

$$N_0 = \sum_{n_s=1}^{N_s} \left[\text{Poisson}\left(V_p \cdot 10^{C_p}\right) + \text{Poisson}\left(V_f \cdot 10^{C_f}\right) \right] \qquad (2)$$

$$N_s = \text{Binomial}\left(N_f, P_p \cdot P_s\right) \qquad (3)$$

The concentration of *C. burnetii* in placentas and foetal fluid was determined from previous studies. Data from natural infection and normal parturition, where the detected *C. burnetii* were shown to be infective by inoculation of rodent models, were used. Placental size and foetal fluid volume were taken from published reports and combined empirically, and single births were assumed.

Exposure assessment

Multiple input parameters were used to simulate the absolute bacterial dose received (table 2). These are described in detail elsewhere (Dewé & Stärk, 2012), except where otherwise indicated. For farmers, exposure was based on the probability of having direct contact (*PD*) with infected material during the reproductive season, in which case an overwhelming dose was assumed. When direct contact did not occur, the exposure dose was assumed equivalent to that of a person 100 m away.

Table 2. Input parameters for exposure assessment[a]

	Definition	Parameterisation	Value
N_0	Total *C. burnetii* shed by infected flock	Output from release assessment	
t	Random day after parturition	Discrete Uniform (min, max)	Min: 0, max: 92
DRT	Time for *C. burnetii* load to reduce to one tenth (days)	Uniform (min, max)	Min:12.7, max: 126 (Roest et al., 2011)
F	Daily aerosolization fraction	Fraction of bacteria in solid waste emitted during one day	0.003
PD	Probability direct contact with infective material occurs	$P_{in} \cdot PD_{in} + P_{out} \cdot PD_{out}$	
P_{in}	Probability parturition occurs inside	Uniform (min, max), where min: Beta (w +1, n – w +1), max: Beta (2-s, n+s), w: number of births in winter, s: number of births in summer, n: annual number of births	Sheep w:13149, s: 812, n: 42978; goat w: 9973, s: 318, n: 18908

	Definition	Parameterisation	Value
P_{out}	Probability parturition occurs outside	1-P_{in}	
PD_{in}	Probability farmer has direct contact with infective material inside		1
PD_{out}	Probability farmer has direct contact with infective material outside	Uniform (min, max)	Min: 0, max: 0.1
Q	Emission rate (*C. burnetii*/s)	$(N_t \cdot F)/86400$	
u	Average wind speed (m/s)	Cumulative distribution	Meteorological data (MeteoSchweiz, 2012)
σ_y	Lateral dispersion coefficient	$4.230+0.9222\ln\left(\frac{x}{1000}\right)-0.0087\ln\left(\frac{x}{1000}\right)^2$	
σ_z	Vertical dispersion coefficient	$3.414+0.7371\ln\left(\frac{x}{1000}\right)-0.0316\ln\left(\frac{x}{1000}\right)^2$	
h	Height of source (m)		2
v	Deposition velocity (m/s)	Uniform (min, max)	Min: 0.001, max: 0.005
x	Distance from infected flock (m)	Pre-defined scenarios	100, 1000, 2000, 3000, 4000 or 5000
z	Height of receptor (m)	Height of average adult	1.70 (Cavelaars et al., 2000)
B	Inhalation volume per hour (m^3/hr)	Inhalation volume of average adult	0.83
T	Exposure time (hours)	Time spent by adult in one place during working day	8

[a] Key: min, minimum value; max, maximum value

Off-farm exposure was modelled by approximating the source strength, or total number of organisms present (N_t), at random time t days after release at parturition (Eq. (4)). Simultaneous parturition within the flock was assumed, and t was restricted to a three-month period, as per a typical Swiss management cycle (S. Thomann, 2012, pers. comm.). The number of *C. burnetii* emitted per second from infected material (Q) was calculated from N_t.

$$N_t = N_0 \cdot 10^{\frac{-t}{DRT}} \qquad (4)$$

Airborne spread of the bacteria was then simulated using a Gaussian plume model. The flock was treated as a single point source and neutral weather conditions were assumed (Sales-Ortells & Medema, 2012). The Gaussian equation for particulate exposure along the centre-line of the plume, at distance x metres from the source is (Lighthart & Frisch, 1976; Lighthart & Mohr, 1987):

$$\overline{C} = \frac{Q}{2\pi u \sigma_y \sigma_z} \exp\left(\frac{-\left[z - \left(h - \frac{vx}{u} \right) \right]^2}{2\sigma_z^2} \right) \tag{5}$$

$\overline{C}$ is the average concentration of *C. burnetii* at x (organisms/m^3). The Pasquill-Gifford dispersion coefficients σ_y and σ_z represent dispersion in the lateral and vertical planes respectively, and are calculated from the McMullen equations (1975). Total exposure dose *ED* (number of organisms) was simulated at each distance from the shedding flock:

$$ED = \overline{C} \cdot B \cdot T \tag{6}$$

Consequence assessment and risk estimation

The consequence assessment was combined with the final risk estimation, that is, the probability of acute symptomatic Q fever in small ruminant farmers during a reproductive season, or in the average Swiss adult on a random day after parturition in a nearby flock. Inputs are shown in Table 3 and described in detail elsewhere (Dewé & Stärk, 2012).

Table 3. Input parameters for consequence assessment[a]

	Definition	Parameterization	Value
ED	Exposure dose of *C. burnetii*	Output from exposure assessment	
P_q	Probability of infection, given susceptibility	$1 - \exp(-ED \cdot \kappa)$	
κ	Dose-response constant	Uniform (min, max)	Min: 0.01, max: 0.9
P_{sus}	Probability a person is susceptible	1-*Sero*	
Sero	Seroprevalence	Pert (min, most likely, max)	General population min: 0.02, most likely: Uniform (0.02,0.1), max: 0.4; farmer min: 0.03, most likely: 0.3, max: 0.7
P_{sym}	Probability of symptoms given infection	Pert (min, most likely, max)	Min: 0.4, most likely: 0.4, max: 0.67
N_F	Number of farmers in Switzerland	Number of flocks	Sheep: 13304, goat: 6822
N_Q	Number of farmers with acute Q fever		

[a] Key: min, minimum value; max, maximum value

The probability of acute symptomatic Q fever (P_{acute}) was estimated as:

$$P_{acute} = P_i \cdot P_f \cdot P_q \cdot P_{sus} \cdot P_{sym} \tag{7}$$

P_q is the probability of infection given that a person is susceptible, and was determined using a single-hit dose-response curve, where one organism has the potential to cause infection and κ is the probability of this occurring. Seroprevalence was used as a proxy for immunity and was a best-bet estimation based on published reports from Switzerland (Dupuis et al., 1986) and Europe (EFSA 2010). The remaining probabilities were drawn from the literature. The number of farmers in Switzerland experiencing acute Q fever was also simulated:

$$N_Q = \text{Binomial}\ (N_F, P_{acute}) \tag{8}$$

Model validation

In a stable population, seroprevalence ($Sero_{app}$) can be approximated by the product of the incidence rate (*IR)* and the duration of antibody titre (*Ab*), described in Table 4. In this model:

$$IR = P_i \cdot P_f \cdot P_q \cdot P_{sus} \tag{9}$$

The Q fever seroprevalence of sheep and goat farmers was simulated separately and then multiplied by the respective proportions of sheep or goat flocks to estimate the prevalence in small ruminant farmers as a whole. This was then compared to the reported seroprevalence for farmers in Europe (EFSA, 2010).

For off-farm scenarios, the ratio (R_x) of the mean probability of symptomatic infection at each distance from a shedding flock, compared to the probability at 5000 metres away, was also simulated. This was compared to the relative risk ratios for symptomatic infection reported during actual outbreaks (Brouqui et al., 2004; Schimmer et al., 2010). Additionally, the simulated attack rate (*AR*) for acute symptomatic infection was calculated over a range of distances from the flock (Eq. (10)), and compared to attack rates observed in the field (Lyytikaeinen et al., 1998; Porten et al., 2006; Gilsdorf et al., 2008). N_{req} is the number of people required for one symptomatic infection to occur.

$$AR = 1/N_{req} \tag{10}$$

Table 4. Input parameters for model validation and sensitivity analysis[a]

	Definition	Parameterization	Value
Ab	Duration of positive antibody titre (years)	Uniform (min, max)	Min: 5, max: 10 (Waag et al., 1995)
R_x	Ratio of P_{acute} at distance x, compared to n metres from a shedding flock	$\frac{P_{acute_x}}{P_{acute_n}}$	x: as before, n: 5000
a	Shape parameter for beta-Poisson curve	As per immuno-competent mice	0.356
N_{50}	Dose at which infection occurs in 50% of exposed individuals	Uniform (min, max)	Min:1, max: 50

[a] Key: min, minimum value; max, maximum value

Sensitivity analysis

Advanced sensitivity analysis was performed in @RISK. Automatic regression analysis identified the most influential parameters on the probability of acute symptomatic infection. Repeated simulations were automatically performed for these parameters, at different percentile values of distributions or at a different percentage change from the base value. The impact of increasing the exposure time ten-fold was assessed, and the approximate beta-Poisson curve was also tested (Table 4), as described elsewhere (Dewé & Stärk, 2012):

$$P_q = 1 - \left[1 + ED \cdot \left(\frac{2^{\frac{1}{a}} - 1}{N_{50}} \right) \right]^{-a} \tag{11}$$

RESULTS

Release assessment

The mean probability of any flock of sheep or goats in Switzerland shedding during a reproductive season was 0.031 [95% CI: 0-0.097] and 0.084 [95% CI: 0.002-0.187], respectively (Fig. 2). The estimated mean number of *C. burnetii* shed per flock was estimated to be 3.93 x 10^{12} organisms [95% CI: 4.66 x 10^{4} – 3.34 x 10^{13}] and 2.93 x 10^{12} [95% CI: 3.42 x 10^{4} – 2.44 x 10^{13}] for sheep and goats, respectively.

Exposure assessment

The mean dose of *C. burnetii* received by a farmer was very high and the mean exposure dose of the average adult off-farm declined with increasing distance from the flock. The absolute exposure dose at each distance was comparable for sheep and goats (not shown).

Consequence assessment and risk estimation

The mean probability of acute symptomatic infection when managing a shedding flock was 0.255 [95% CI: 0-0.415] for sheep farmers; results were similar for goat farmers. For all Switzerland, these probabilities were 0.008 [95%CI: 0-0.031] and 0.022 [95%CI: 0-0.061], respectively. This translated to mean numbers of 110 [95% CI: 0-423] acute Q fever cases in sheep farmers and 194 [95% CI: 0-535] in goat farmers, annually.

Off-farm, probability of symptomatic infection ranged from 0.131 [95% CI: 0-0.457] for a typical adult at 100 metres away from a shedding sheep flock, to 0.023 [95% CI: 0-0.338] at five kilometres. Results were similar for goat flocks, although for the country as a whole, goat flocks were shown to represent the greater risk. For example, the probability of symptomatic infection of a typical adult one kilometre away from a sheep flock in Switzerland was 0.003 [95%CI: 0-0.027], and from a goat flock, 0.006 [95%CI: 0-0.054] (Fig. 2).

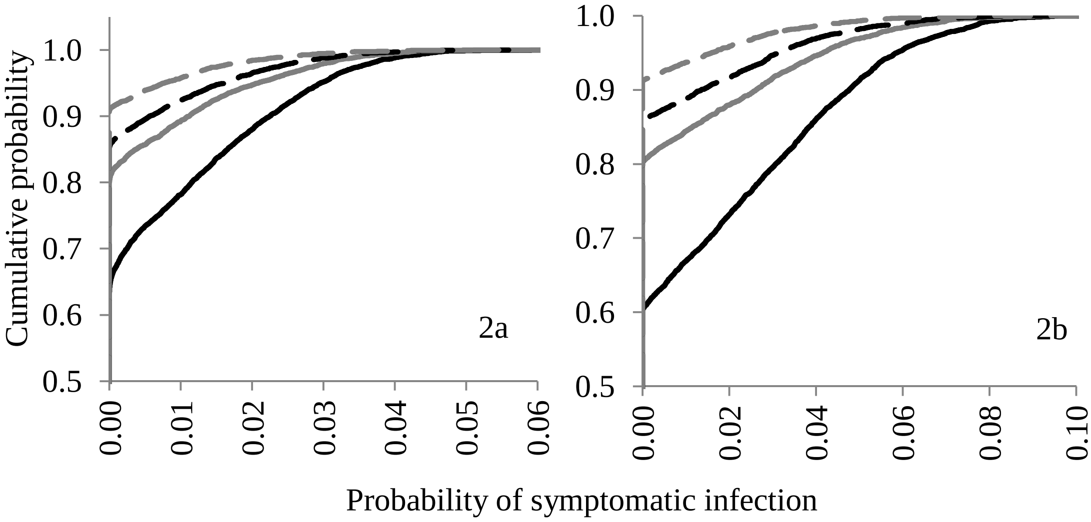


Fig. 2 Cumulative distribution of the probability of acute symptomatic *Coxiella burnetii* infection in an average adult downwind from a sheep (2a) or goat (2b) flock in Switzerland. Distance from flock: 0.1 km (——); 1.0 km, (——); 2.0 km (– – –); 5.0 km (– – .)

Model validation

The simulated risk was greatest for farmers, and declined over increasing distance from a shedding flock, which is consistent with known risk factors (Veenstra et al., 2011). The mean Q fever seroprevalence in small ruminant farmers was calculated to be 0.235 [95% CI: 0.005-0.625], which is within the range reported for this group (EFSA, 2010).

The simulated relative risk of acute infection off-farm declined with increasing distance from the flock, consistent with the relative risk observed during actual outbreaks (Fig. 3a), and was similar for sheep and goat flocks. Acute symptomatic infection of people off-farm occurred in 37.5% of simulations at 100 metres away from a shedding sheep flock, and 9.3% of simulations at five kilometres away. The attack rate on such days was comparable to attack rates reported during actual outbreaks, and results from goat flocks were similar.

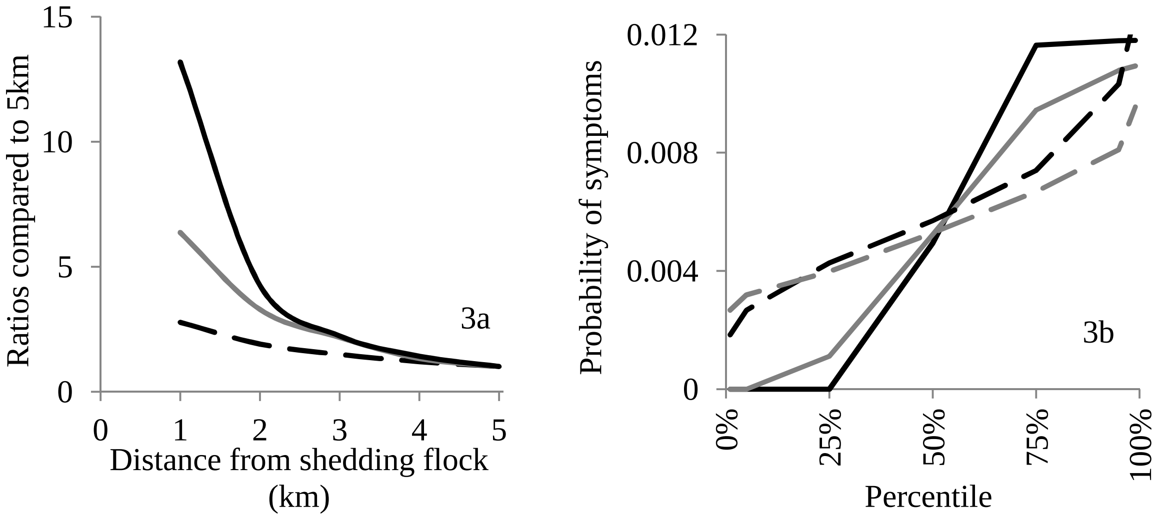


Fig. 3a Results of *C. burnetii* risk assessment compared with literature. Ratios of simulated exposure dose (▬) and probability of symptomatic infection (▬ ▬); risk ratios of symptomatic infection observed in Schimmer et al. (2010) (▬). Fig. 3b Sensitivity analysis showing the change in mean probability of acute symptomatic *C. burnetii* infection 1.0 km downwind from a Swiss goat flock, expressed as percentiles of the most influential input distributions. Concentration of *C. burnetii in* placenta (▬); within-flock shedding prevalence (▬); flock infection prevalence (▬ ▬); flock size (▬ ▬)

Sensitivity analysis

Sensitivity analysis showed that the probability of symptomatic infection (P_{acute}) was most sensitive to the *C. burnetii* concentration in placenta (Fig. 4b). Other influential inputs were the within-flock shedding prevalence, flock infection prevalence, and flock size. *C. burnetii* concentration in foetal fluids had a lesser effect and seroprevalence had a mild negative influence (not shown). The model was minimally sensitive to the probability that one organism causes infection in the human host (κ). However, test simulations using the beta-Poisson dose response curve reduced the simulated probability of symptomatic infection of a farmer by approximately one third (results not shown). A greater difference was seen at lower exposure doses (i.e. increased distance from the flock). Increasing the exposure time ten-fold increased $P_{acute\ by}$ approximately one sixth (results not shown).

DISCUSSION

This study is the first to develop a quantitative, stochastic approach for assessing the risk of livestock-associated human Q fever infection. Parameters used in this model include the flock seroprevalence, probability of shedding within a flock, the quantities of *C. burnetii* shed during parturition, the dispersion of airborne particles over increasing distance from a flock, an estimation of the *C. burnetii*-human dose-response relationship, and human susceptibility to infection. A number of uncertainties were identified during this process; however, we are confident that the results demonstrate a non-negligible risk of Q fever in Switzerland.

Release assessment

The risk of *C. burnetii* release from sheep and goat flocks was modelled for the first time. Release inputs with greatest influence on the overall risk to public health in this model are the concentration of *C. burnetii* in placenta, the prevalence of flock infection and within-flock shedding, and the flock size.

To the authors' knowledge, the distribution of the number of *C. burnetii* shed during an infective parturition has not been previously estimated. The values used here were drawn from early studies but were considered the most representative. More recent work tends to report shedding from experimentally infected animals or during abortion investigations, and almost exclusively involve the use of PCR techniques, which do not demonstrate bacterial viability. The concentration of *C. burnetii* in goat placenta and birth fluids was extrapolated from sheep data, as limited information is available for this species. This could result in an under-estimation of the differences in risk represented by the two species. Multiple births and breed variation were not considered here, although these are likely to have less impact on *C. burnetii* release.

A key parameter of the model is the proportion of infected flocks, for which seroprevalence was used as a proxy. This approach has limitations: the presence of antibodies does not distinguish between current and past infections, and flock seronegativity does not exclude the possibility of shedding at parturition; however, seropositivity was considered the best approximation of flock infection from the data available.

The relationship between within-flock seroprevalence and within-flock shedding was also explored here for the first time, and no association could be found. Instead, within-flock shedding prevalence was modelled by an empirical distribution based on published data from seropositive flocks. This probably resulted in an over-estimation of the overall risk, as the majority of studies involved flocks with a known history of *C. burnetii* infection or reproductive problems. Shedding is perhaps more prevalent amongst primiparous animals (Hogerwerf et al., 2011), but this was not explored here. Average within-flock shedding prevalence was estimated to be 0.257 [95% CI: 0-0.944] and 0.388 [95%CI: 0.005-0.947], for sheep and goats respectively, although the most likely value was close to zero.

The number of animals shedding within a flock is dependent on shedding prevalence and on flock size, both of which were critical to determining the overall risk to the general public. Risk to farmers, however, is not sensitive to the number of shedding animals if this is non-zero, due to the large exposure dose received during close contact with infective material from a single parturition. Small flocks were under-represented in the serological survey, compared to the population, although the effect of flock size on within-flock shedding prevalence is unknown.

The dynamics of within- and between-flock transmission and the potential for abortion storms were not explored here. In this model, parturition within a flock was simulated to occur within the same day, which is perhaps temporally more representative of abortions than natural parturitions. Abortions could therefore increase the simulated public health risk if aborted material contains higher concentrations of *C. burnetii,* which is not necessarily the case (Roest et al., 2012), or if they significantly alter the within- or between-flock shedding prevalence. If so, goats could present an additional risk, as *C. burnetii* abortion in Switzerland has been shown to occur more frequently in this species (Chanton-Greutmann et al., 2002).

Shedding in the vaginal mucus and milk of small ruminants has been detected previously (Arricau-Bouvery et al., 2003). However, these release pathways were not modelled in this study after preliminary sensitivity analysis showed that they contributed minimally to environmental contamination and the overall risk (data not shown). Faecal shedding was also excluded from the analysis, as very little evidence exists to support intestinal excretion of viable *C. burnetii*, rather than passive transport through the gut following ingestion of contaminated material (Abinanti et al., 1953; Winn et al., 1961; Roest et al., 2012).

Exposure assessment

The exposure dose of *C. burnetii* received by farmers and individuals at a range of distances from a shedding flock was heavily influenced by the release inputs, and only slightly sensitive to the exposure parameters of this model.

The *C. burnetii* source strength could be overestimated by modelling all parturitions as occurring on the same day, although this is unlikely to have a significant effect on the overall risk, due to the large numbers of bacteria that can be shed during a single parturition. Other assumptions were made due to data gaps, particularly regarding the parameters affecting the rates of *C. burnetii* decay and aerosolization. While *C. burnetii* is known to persist in the environment for a long time (Maurin & Raoult, 1999), death rates have not been extensively investigated for this bacteria, and the *DRT* used here is an extrapolation of experimental work at higher temperatures (Roest et al., 2011). Additionally, the aerosolization factor is an estimation that, while usually applied to a range of bacterial species (Dungan, 2010), is not specific to *C. burnetii.*

Gaussian plume dispersion models are most suited to flat rather than complex (i.e. hilly or mountainous) terrain (EPA, 1993). However, they are widely accepted in the literature, are relatively simple to compute, and have been previously used as screening tools to assess *C. burnetii* dispersion (Sales-Ortells & Medema, 2012). While alternative dispersion models exist, these were either unsuitable for modelling *C. burnetii* release, or required detailed meteorological data, thus making them inapplicable to the generic situation considered here. Multiple puff-particle models (e.g. CALPUFF, TRC Solutions, Lowell, USA) could, however, be utilised for site-specific risk assessment in the future.

The deposition velocity incorporated in the plume model is dependent on the size of the particles being transported. While *C. burnetii* itself is less than 1 μm diameter (Maurin & Raoult, 1999) , it has been detected in particulate matter of 2.5-10.0 μm diameter (Hogerwerf et al., 2012). Although not explored extensively, preliminary work indicates that *C. burnetii* in goat houses might be relatively evenly distributed across this range of particle sizes (Aarnink et al., 2012).

Off-farm exposure was modelled for a random day during a three-month period after parturition. This was used as a crude proxy for the as-yet undefined combination of meteorological and geographic factors that could be required for airborne *C. burnetii* spread (Tissot-Dupont et al., 1999; van der Hoek et al., 2011). Interestingly, sensitivity analysis showed that increasing exposure time had a relatively small effect on the overall probability of acute Q fever (results not shown). This is likely to reflect the relatively high probability of infection given exposure that was used in this model.

In this model, the dose received by a farmer was unaffected by the assumptions made regarding *C. burnetii* emission and dispersion. Close contact with material from a single infective parturition was sufficient to deliver a massive exposure dose, due to the high concentrations of *C. burnetii* in placenta and birth fluids, and probability of symptomatic infection was dependent on the likelihood of this contact occurring. While simplifying assumptions could reduce the accuracy of this estimation, it enabled the annual risk of symptomatic infection to be approximated.

Consequence assessment and risk estimation

The consequence assessment, and overall risk estimation, was based on the *C. burnetii* dose received from a shedding flock, the dose-response relationship, and the probabilities of susceptibility and the development of symptoms. Goat flocks were shown to represent a greater risk to public health than sheep flocks, which is a reflection of an increased probability of shedding in this model. This is attributable both to the increased seroprevalence of Swiss goat flocks, and to the higher within-flock shedding prevalence reported in caprine Q fever investigations.

The model showed farmers to be at greater risk than the general public, for whom the Q fever risk decreased with increasing distance from shedding flocks. Both of these findings are consistent with observations made under natural conditions (Garner et al., 1997; Schimmer et al., 2010) and support the use of this model to approximate symptomatic Q fever over a range of distances. Attack rate over limited distances from a shedding flock is within the range reported during other outbreaks, which further validates the model.

Human susceptibility to infection was approximated by (1 – seroprevalence). Cellular immunity, which has an important role in protection against Q fever infection (Maurin & Raoult, 1999), was not considered here, and so susceptibility could be overestimated. However, serological results are much more commonly reported and were considered suitable for purpose.

The probability of infection is determined by the *C. burnetii* dose-response curve, which has not yet been determined for humans; however, it has been explored for guinea pigs following intra-peritoneal inoculation (Jones et al., 2006). The authors assumed that guinea-pigs and humans are equally susceptible to *C. burnetii* (Tigertt et al., 1961), but that the probability of infection is higher for intraperitoneal inoculation than for inhalation under field conditions. Sensitivity analysis showed that κ had a minor influence on the overall risk; this was affected more by using the alternative beta-Poisson curve.

The probability of people developing clinical signs once infection has occurred has been widely reported as 40% (EFSA, 2010), although variations have been recorded. Q fever has rarely been reported in Switzerland, but this study estimates that on average, approximately 300 small ruminant farmers could be affected each year. Differences could be due to overestimation of the overall risk by the model, or to under-reporting, particularly as the disease is non-notifiable.

Conclusions and recommendations

Under endemic, steady-state conditions, approximately 3% and 8% of sheep and goat flocks in Switzerland can be expected to shed *C. burnetii* during a reproductive season. When shedding occurs, massive numbers of *C. burnetii* - in the order of 10^{12} organisms on average -

can be released from a single flock. According to this model, goat flocks present a greater risk than sheep flocks. Farmers were shown to be at greater risk of symptomatic Q fever infection than the general public, for whom the risk was shown to decrease with increasing distance from the animals. On average, approximately 300 small ruminant farmers could be experiencing acute Q fever symptoms in Switzerland each year.

Appropriate risk management at this stage could involve increasing awareness of Q fever amongst small ruminant farmers, medical personnel, and rural populations in general, until at-risk individuals or regions can be identified and targeted appropriately. It is hoped that the original risk assessment framework developed here can inform on further research that could reduce the uncertainty of input parameters, thereby allowing more accurate quantification of the public health risk, and ultimately, the effect of control measures. Suggested priorities for veterinary public health professionals are determining the amount of viable *C. burnetii* released during normal parturition and abortion more accurately, and establishing the between- and within-flock prevalence of *C. burnetii* shedding.

ACKNOWLEDGEMENTS

We acknowledge the suggestion of a technical improvement to the model by A. Hill, and are grateful to A. Lindberg for her editorial input and to H.-J. Roest for his generous discussions. We would also like to thank H.-P. Schwermer, S. Thomann and C. Aeschlimann for providing data, and M. Schuppers for reviewing the original manuscript. The project was funded by the Swiss Federal Veterinary Office, Bern, Switzerland, and completed in part fulfilment of an MSc in Control of Infectious Diseases at the Royal Veterinary College, London, United Kingdom. TD was supported by the W.E.J. Craig Prestigious Travelling Scholarship from the University of Melbourne, Victoria, Australia.

REFERENCES

Aarnink, A.J.A., Roest, H.-J., Cambra-Lopez, M., Zhao, Y., Mosquera, J. and Ogink, N.W.M. (2012). Emissions and concentrations of dust and pathogens from goat houses. Int. Livest. Environ. Symp. 9, paper no. ILES12-0968

Abinanti, F.R., Lennette, E.H., Winn, J.F. and Welsh, H.H. (1953). Q fever studies: XVIII. Presence of *Coxiella burnetii* in the birth fluids of naturally infected sheep. Am. J. Hyg. 58, 385-388

Aeschlimann, C. (2012) Sheep stud book data. Schweizerischer Schafzuchtverband SZV, Bern, Switzerland

Arricau-Bouvery, N., Souriau, A., Lechopier, P. and Rodolakis, A. (2003). Experimental *Coxiella burnetii* infection in pregnant goats: excretion routes. Vet Res 34, 423-433

Brouqui, P., Badiaga, S. and Raoult, D. (2004). Q fever outbreak in a homeless shelter. Emerg. Infect. Dis. 10, 1297-1299

Cavelaars, A.E.J.M., Kunst, A.E., Geurts, J.J.M., Crialesi, R., Grötvedt, L., Helmert, U., Lahelma, E., Lundberg, O., Mielck, A., Rasmussion, N.K., Regidor, E., Spuhler, T. and

Mackenbacch, P. Persistent variations in average height between countries and between socio-economic groups: an overview of 10 European countries. Ann. Hum. Biol. 27, 407-421

Chanton-Greutmann, H., Thoma, R., Corboz, L., Borel, N. and Pospischil, A. (2002). [Abortion in small ruminants in Switzerland: investigation during two lambing seasons (1996-1998) with special regard to chlamydial abortions]. Schweiz. Med. Wschr. 144, 483-492

Dewé, T.C.M. and Stärk, K.D.C. (2012). *Coxiella burnetii* in small ruminants in Switzerland: assessing the public health risk. Swiss Federal Veterinary Office, Bern, Switzerland. http://www.aramis.admin.ch/Dokument.aspx?DocumentID=2675 [Accessed online 23 January 2013]

Dungan, R.S. (2010). Fate and transport of bioaerosols associated with livestock operations and manures. J. Anim. Sci. 88, 3693-3706

Dupuis, G., Péter, O., Mottiez, M.C. and Vouilloz, M. (1986). Séro-prévalence de la fièvre Q humaine en Suisse. Schweiz. Med. Wschr. 116, 494-498

Dupuis, G., Petite, J., Péter, O. and Vouilloz, M. (1987). [An important outbreak of human Q fever in a Swiss alpine valley]. Int. J. Epidemiol. 16, 282-287

ECDC. (2010). ECDC technical report: risk assessment on Q fever. European Centre for Disease Prevention and Control, Stockholm, Sweden

EFSA. (2010). Scientific opinion on Q fever. EFSA J. 8, 1595-1709

EPA. (1992). Guidelines for exposure assessment. US Environmental Protection Agency, Washington D.C., USA

EPA. (1993). Selection criteria for mathematical models used in exposure assessments: atmospheric dispersion models. US Environmental Protection Agency, Washington D.C., USA

Garner, M.G., Longbottom, H.M., Cannon, R.M. and Plant, A.J. (1997). A review of Q fever in Australia 1991-1994. Aust. N.Z. J. Publ. Heal. 21, 722-730

Gilsdorf, A., Kroh, C., Grimm, S., Jensen, E., Wagner-Wiening, C., and Alpers, K. (2008). Large Q fever outbreak due to sheep farming near residential areas, Germany, 2005. Epidemiol. Infect. 136, 1084–1087

Hawker, J.I., Ayres, J.G., Blair, I., Evans, M.R., Smith, D.L., Smith, E.G., Burge, P.S., Carpenter, M.J., Caul, E.O., Coupland, B., Desselberger, U., Farrell, I.D., Saunders, P.J. and Wood, M.J. (1998). A large outbreak of Q fever in the West Midlands: windborne spread into a metropolitan area? Comm. Dis. Publ. Heal. 1, 180-187

Hilbert, A., Schmoock, G., Lenzko, H., Moog, U., Diller, R., Fröhlich, A., Hoffmann, L., Horner, S., Elschner, M., Tomaso, H., Henning, K., Neubauer, H. and Sprague, L.D.

(2012). Prevalence of *Coxiella burnetii* in clinically healthy German sheep flocks. BMC Res. Notes 5, 152

Hogerwerf, L., van den Brom, R., Roest, H.-J., Bouma, A., Vellema, P., Pieterse, M., Dercksen, D. and Nielen, M. (2011). Reduction of *Coxiella burnetii* prevalence by vaccination of goats and sheep in the Netherlands. Emerg. Infect. Dis. 17, 379-386

Hogerwerf, L., Borlee, .F, Still, K., Heederik, D., Van Rotterdam, B., De Bruin, A., Nielen, M. and Wouters, I.M. (2012). Detection of *Coxiella burnetii* DNA in inhalable airborne dust samples from goat farms after mandatory culling. Appl. Environ. Microbiol. 78, 5410-5412

Hunninghaus, J., Scherrer, S., Schüpbach-Regula, G., Wittenbrink, M.M., Hamburger, A., Stärk, K.D.C. and Magouras, I. (2012). Prevalence of Q fever in small ruminants in Switzerland – implications for public health. Swiss Federal Veterinary Office, Bern, Switzerland. http://www.aramis.admin.ch/Dokument.aspx?DocumentID=2676 [Accessed online 23 January 2013]

Jones, R.M., Nicas, M., Hubbard, A.E. and Reingold, A.L. (2006). The infectious dose of *Coxiella burnetii* (Q fever). Appl. Biosaf. 11, 32-41

Lighthart, B. and Frisch, M. (1976). Estimation of viable airborne microbes downwind from a point source. Appl.Environ.Microbiol. 31, 700-704

Lighthart, B. and Mohr, A.J. (1987). Estimating downwind concentrations of viable airborne microorganisms in dynamic atmospheric conditions. Appl. Environ. Microbiol 53, 1580-1583

Lyytikaeinen, O., Ziese, T., Schwartlaender, B., Matzdorff, P., Kuhnhen, C., Jaeger, C. and Petersen, L. (1998). An outbreak of sheep-associated Q fever in a rural community in Germany. Eur. J. Epidemiol. 14, 193-199

Marrie, T.J. (1995). *Coxiella burnetii* (Q fever) pneumonia. Clin. Infect. Dis. 21, S253- S264

Maurin, M. and Raoult, D. (1999). Q fever. Clin.Microbiol.Rev. 12, 518-553

McMullen, R.W. (1975). The change of concentration standard deviations with distance. JAPCA J. Air Waste MA. 25, 1057-1058

Meteoschweiz. (2012). Langjährige Mittelwerte 1981-2000: Windgeschwindigkeit skalar. Swiss Department for Meterology and Climate, Zürich, Switzerland. www.meteoschweiz.ch [Accessed online 17th June 2012]

OIE. (2012). Terrestrial Animal Health Code. World Organisation for Animal Health, Paris, France. http://www.oie.int/international-standard-setting/terrestrial-code/access-online/ [Accessed online 23rd January 2013]

Porten, K., Rissland, J., Tigges, A., Broll, S., Hopp, W., Lunemann, M., van Treeck, U., Kimmig, P., Brockmann, S., Wagner-Wiening, C., Hellenbrand, W. and Buchholz, U.

(2006). A super-spreading ewe infects hundreds with Q fever at a farmers' market in Germany. BMC Infect. Dis. 6, 147

Roest, H.-J., Tilburg, J.J.H.C., Van Der Hoek, W., Vellema, P., Van Zijderveld, F.G., Klaassen, C.H.W. and Raoult, D. (2010). The Q fever epidemic in The Netherlands: history, onset, response and reflection. Epidemiol. Infect. 139, 1-12

Roest, H.-J., Dinkla, A., van Rotterdam, B., de Bruijn, A., Dercksen, D., and Vellema, P. (2011). [Survival of *Coxiella burnetii* in goat manure]. Dutch Animal Health Service (GD), National Institute for Public Health and the Environment (RIVM) and the Central Veterinary Institute, (CVI), The Netherlands. http://library.wur.nl/WebQuery/clc/1970269 [Accessed online 23rd June 2012]

Roest, H.-J., van Gelderen, B., Dinkla, A., Frangoulidis, D., van Zijderveld, F., Rebel, J., and van Keulen, L. (2012). Q fever in pregnant goats: pathogenesis and excretion of *Coxiella burnetii.* PLoS ONE 7, e48949

Sales-Ortells, H. and Medema, G. (2012). Screening-level risk assessment of *Coxiella burnetii* (Q fever) transmission via aeration of drinking water. Environ. Sci. Tech. 46, 4125-4133

Schimmer, B., Ter Schegget, R., Wegdam, M., Zuchner, L., De Bruin, A., Schneeberger, P.M., Veenstra, T., Vellema, P. and Van Der Hoek, W. (2010). The use of a geographic information system to identify a dairy goat farm as the most likely source of an urban Q-fever outbreak. BMC Infect. Dis. 10, 69-76

Schwermer, H-P. (2012). Agricultural census data. Swiss Federal Veterinary Office, Bern, Switzerland

Tigertt, W.D., Benenson, A.S. and Gochenour, W.S. (1961). Airborne Q fever. Bact. Rev. 25, 285-293

Thomann, S. (2012). Goat stud book data. Schweizerischer Ziegenzuchtverband SZZV, Bern, Switzerland

Tissot-Dupont, H., Torres, S., Meyer, N. and Raoult, D. (1999). Hyperendemic focus of Q fever related to sheep and wind. Am. J. Epidemiol. 150, 67-74

Van Der Hoek, W., Hunink, J., Vellema, P. and Droogers, P. (2011). Q fever in The Netherlands: the role of local environmental conditions. Int. J. Environ. Heal. Res. 21, 441-451

Veenstra, T., Snijders, B.E., Schimmer, B., Rietveld, A., Van Dam, S., Schneeberger, P.M., Dijkstra, F., Van Der Sande, M.A. and Van Der Hoek, W. (2011). Risk factors for Q fever in the Netherlands. Webmed Central Infect. Dis. 2, WMC002006

Waag, DM. (2007). *Coxiella burnetii*: host and bacterial responses to infection. Vaccine 25, 7288-7295

Welsh, H.H., Lennette, E.H., Abinanti, F.R. and Winn, J.F. (1958). Air-borne transmission of Q fever: the role of parturition in the generation of infective aerosols. Ann. N.Y. Acad. Sci. 70, 528-540

Winn, J.F., Abinanti, F.R., Lennette, E.H and Welsh, H.H. (1961). Q fever studies: XXII. Inoculation of sheep by the intestinal route. Am. J. Hyg. 73, 105-113

DETERMINANTS FOR PERSISTENT LIVESTOCK-ASSOCIATED MRSA CARRIAGE IN VEAL CALF FARMERS AND THEIR FAMILY MEMBERS

A. DORADO-GARCIA[*], M.E.H. BOS, H. GRAVELAND, B.A.G.L. VAN CLEEF,
K.M. VERSTAPPEN, J. KLUYTMANS, J.A. WAGENAAR AND D.J.J. HEEDERIK

SUMMARY

A longitudinal study was performed over a period of one week in 51 randomly selected veal calf farms to assess risk factors for persistent LA-MRSA nasal carriage among farmers and household members. On day 0, swabs were taken from nose and throat of participants and on days 4 and 7 swabs were taken only from the nose. A persistent MRSA carrier was defined as a person whose nasal swabs were positive for MRSA presence on all days. A questionnaire was used to identify potential risk factors. Data were analysed using Generalized Estimating Equations (GEE) modelling and Poisson regression. The outcome was positively associated with contact with veal calves, administration of antimicrobials to animals and the presence of potential animal reservoirs. Furthermore, MRSA persistent carriers were characterised by higher MRSA CFU counts, presence of MRSA in throat and absence of MSSA.

INTRODUCTION

In recent years, livestock-associated methicillin-resistant *Staphylococcus aureus* (LA-MRSA), specifically sequence type (ST) 398, has emerged in food-producing animals and people in contact with these animals (van Duijkeren et al., 2004, Graveland et al., 2011a, Fitzgerald, 2012). Illness associated with ST398 in humans is rare and only a small proportion of MRSA infections can be attributed to LA-MRSA (van Cleef et al., 2011). Nonetheless, invasive infections and hospital outbreaks of MRSA ST398 have been reported in Europe, the United States and Asia (Rasigade et al., 2010, van Cleef et al., 2011, van der Mee-Marquet et al., 2011).

LA-MRSA strains have been found mainly in pigs and veal calves, but they have the capacity to colonise a wide spectrum of hosts, including sheep and poultry. However, the emergence of LA-MRSA in humans and the possibility of having new *mecA* gene alleles circulating in the environment are major causes of concern since it can give rise to new and more virulent MRSA strains (Fitzgerald, 2012). The potential public health threat posed by LA-MRSA carriage in humans justifies the demand for more knowledge about the epidemiology of these strains.

---

[*] Alejandro Dorado García, Institute for Risk Assessment Sciences, Utrecht University, PO Box 80178, 3508 TD Utrecht, The Netherlands. Email: a.doradogarcia@uu.nl

A longitudinal study including periods of high and low exposure to animals showed that LA-MRSA carriage was transient and it was suggested that LA-MRSA is a poor persistent coloniser in humans (Graveland et al., 2011b). However, risk factors for LA-MRSA persistent carriage and for a possible true colonisation have not been thoroughly assessed.

The current study evaluates risk factors for persistent MRSA carriage over a period of one week in veal calf farmers and their family members. The study focuses mainly on occupational practices, business management and household characteristics among other potential determinants.

MATERIALS AND METHODS

Study design

A longitudinal study was performed over a period of one week at the beginning of the veal production cycle in 51 randomly selected veal calf farms. All farms were visited from the end of 2010 to the end of 2011. On each farm nasal swabs from both anterior nares were taken and analysed in 10 pools of 6 calves each (60 animals per farm) at day 0. Swabs were also collected from farmers, family members and employees. On day 0 quantitative swabs were taken by field workers from the nose (both nostrils) and throat in the majority of participants. On days 4 and 7 dry swabs were used to self-sample only the nose. The swabs were immediately shipped or mailed to the laboratory and processed within 24 hours after arrival. For the purpose of this study, a person was defined to be a persistent MRSA carrier when each of the nasal swabs collected on days 0, 4 and 7 was positive for MRSA presence.

All participants gave their informed consent and were asked to fill in an extensive questionnaire which included items on potential risk factors. The protocol of the study was approved by Medical Ethical Committee of Utrecht University. The collection of animal samples was in compliance with the Dutch Law on Animal Health and Welfare.

Laboratory analysis

Swabs were given to participants with instructions including photographs in case of self-sampling. Swabs in liquid transport medium (Eswab, Copan, Brescia, Italy) were used for quantitative cultivation. Serial dilutions were cultured on chromID *S. aureus* and chromID MRSA agar plates (BioMérieux, La Balme Les Grottes, France) at 37°C overnight. Plates with 10-100 CFU were used to calculate the original amount of CFU per swab. In order to detect positive samples without bacterial growth on the first day, the remaining transport medium and swab were enriched overnight in Mueller Hinton broth with 6.5% NaCl (MH), and consequently cultured on chromID *S. aureus* and chromID MRSA agar plates. The theoretical lower limit of quantification (LLOQ) of MRSA colony forming units (CFU) was 10. Dry cotton swabs (Copan, Brescia, Italy) taken on days 4 and 7 were inoculated directly onto chromID *S. aureus*, chromID MRSA and MH. Confirmation of MRSA species on the 3 sampling times was done by Real-Time PCR targeting *mecA*, *femA* and *nuc* genes. MSSA presence was tested when the bacterial growth in chromID *S. aureus* was higher than in chromID MRSA. For this purpose, several colonies were screened for methicillin susceptibility by using the cefoxitin disk diffusion method. Confirmation of methicillin-susceptible *Staphylococcus aureus* (MSSA) species was done by Real-Time PCR as described above.

Nasal swabs from calves were analysed in pools of 6 animals following standard procedures previously described by Graveland and co-workers (2010).

Data analysis

Statistical analysis was performed using SAS software version 9.2 (SAS institute Inc., Cary, North Carolina, USA).

Crude prevalence ratios (PRs) and ratios adjusted for sex and age were obtained for nasal MRSA persistent carriage. PROC GENMOD was used for Generalized Estimating Equations (GEE) modelling to take clustering of data at farm level into account. The mean response was modelled with a Poisson regression with robust standard errors. A p-value <0.05 was considered statistically significant.

The number of CFU from quantitative nasal swabs positive to MRSA but below LLOQ was set to 5 for data analysis purposes.

RESULTS

Nasal carriage patterns for MRSA, MSSA and *S. aureus* are given in Table 1. MRSA and MSSA persistent carrier prevalence followed opposite directions in farmers as compared to family members. For MRSA persistent carriage, the prevalence in farmers (15.5%) was twice as high as in family members (7.6%). MSSA persistent carriage prevalence was three times higher in family members as compared to farmers (15.3% and 5.2%, respectively). Regarding *S. aureus*, 22.8% of the individuals were persistently carrying the bacteria, 29.6% were intermittent carriers and the remaining 47.6% never carried *S. aureus.*

No persistent MRSA carrier was positive for MSSA at day 0. People found positive for MRSA in throat swabs at day 0 were at higher risk for being persistent nasal carriers (adjusted PR=12.2, 95% CI=5.2-28.8, p<0.0001).

CFU counts were determined in 42 participants from quantitative nasal swabs on day 0. The median CFU count per swab suspension was 44 with an interquartile range (IQR) 5-1,100. The univariate logistic regression analysis in this population resulted in 1.7 times higher risk (95% CI=1.3-2.1, p<0.001) for persistent MRSA carriage per 10 CFU increase.

Crude and age-sex-adjusted PRs in determinants meeting the specified criteria are presented in Table 2. Sensitivity analysis in which an additional adjustment was made for smoking habits did not result in significant changes in estimates.

Statistically significant risk factors for persistent MRSA carriage were identified (Table 2). These include pet ownership and the presence of cats and sheep on the farm as well as performing specific tasks on the farm. The number of working hours per week on the farm showed the most significant association with the outcome (adjusted PR=2.5 expressed per 20 hours/week increase, p=0.001). Administration of antimicrobials to calves through milk and injection in the past month preceding sampling was also a significant risk factor (adjusted PR= 3.4, p=0.01). Other determinants such as the prevalence of MRSA in animals at the farm level did not show an association with persistent human MRSA carriage (PR=1, 95%CI=1.0-1.0, p=0.96).

Table 1. Patterns of nasal carriage of *S. aureus*, MRSA and MSSA in the study total population and subpopulations of farmers, household members and employees

	No. persistent (%)[a]	No. intermittent (%)[a]	No. non-carrier (%)[a]	Total no.
MRSA in nose:				
Total population	20 (9.7)	35 (17.0)	151 (73.3)	206
Farmers	9 (15.5)	15 (25.9)	34 (58.6)	58
Family members	11 (7.6)	20 (13.9)	113 (78.5)	144
Employees	0 (0.0)	0 (0.0)	4 (100.0)	4
MSSA in nose:				
Total population	25 (12.1)	36 (17.5)	145 (70.4)	206
Farmers	3 (5.2)	13 (22.4)	42 (72.4)	58
Family members	22 (15.3)	22 (15.3)	100 (69.4)	144
Employees	0 (0.0)	1 (25.0)	3 (75.0)	4
S. aureus in nose:				
Total population	47 (22.8)	61 (29.6)	98 (47.6)	206
Farmers	14 (24.1)	21 (36.2)	23 (39.7)	58
Family members	33 (22.9)	39 (27.1)	72 (50.0)	144
Employees	0 (0.0)	1 (25.0)	3 (75.0)	4

[a] A person was persistent carrier when each of the nasal swabs collected on days 0, 4 and 7 was positive for MRSA; non-carriers had no positive cultures; intermittent carriers were the remaining persons.

Table 2. Crude and adjusted for sex and age prevalence ratios (PR) for nasal MRSA persistent carriage in 195 veal calf farmers and household members

Determinant[a]	Category	N	No. persistent carriers[b] (%)	Mean (range)	PR	95% CI	PR[c] Adj	95% CI
General characteristics:								
Sex	Male	103	9 (8.7)	-	1	-	-	-
	Female	92	11 (12.0)	-	1.4	0.6-3.2	-	-
Age	-	195	-	30 (0.1-81)	1.0	1.0-1.0**	-	-
per 10 years increase	-	195	-	-	1.3	1.1-1.6**	-	-
Business and household characteristics:								
Presence of sheep in farm	No	149	12 (8.1)	-	1	-	1	-
	Yes	46	8 (17.4)	-	2.2	1.1-4.5*	2.4	1.2-4.8*
Presence of cats on farm	No	96	5 (5.2)	-	1	-	1	-
	Yes	99	15 (15.2)	-	3.0	1.2-7.1*	2.7	1.1-6.6*
Presence of pets	No	74	4 (5.4)	-	1	-	1	-
	Yes	121	16 (13.2)	-	2.7	1.0-7.4*	2.6	1.0-6.7[t]

Determinant[a]	Category	N	No. persistent carriers[b] (%)	Mean (range)	PR	95% CI	PR[c] Adj	95% CI
Tasks performed last 7 days[d] :								
Sorting calves	No	113	5 (4.4)	-	1	-	1	-
	Yes	82	15 (18.3)	-	4.2	1.5-12.3**	4.7	1.3-16.8*
Healthcare / control [e]	No	132	9 (6.8)	-	1	-	1	-
	Yes	63	11 (17.5)	-	2.6	1.1-6.1*	2.3	0.8-7.3
Feeding calves	No	72	2 (2.8)	-	1	-	1	-
	Yes	123	18 (14.6)	-	7.2	0.9-58.6[t]	5.4	0.6-52.3
Work at farm, hygiene cleaning and disinfection								
Administration of antibiotics by farmer during last month	No	131	8 (6.1)	-	1	-	1	-
	Yes	64	12 (18.8)	-	3.2	1.4-7.1**	3.4	1.3-9.1*
# working hours per week	-	195	-	16.5 (0-80)	1.0	1.0-1.0***	1.0	1.0-1.1**
per 20 hours increase	-	-	-	-	1.8	1.4-2.4***	2.5	1.4-4.2**
Clean towel	No	45	7 (16.7)	-	1	-	1	-
	Yes	150	13 (8.67)	-	0.6	0.3-1.3	0.6	0.3-1.1
Changing room available	No	18	3 (16.7)	-	1	-	1	
	Yes	177	17 (9.7)	-	0.6	0.3-1.2	0.5	0.2-1.0[t]
Cleaning of baby boxes	No	184	18 (9.8)	-	1	-	1	-
	Yes	11	2 (18.2)	-	1.9	1.0-3.5*	1.3	0.6-2.8

[a] Eligibility criteria for variables included univariate p-values below 0.2, less than 10% of missing data in relation with the outcome, and at least 2 persistent carriers falling in each of the categories of the explanatory categorical variables.

[b] A person is considered a persistent carrier when all nasal swabs at days 0, 4 and 7 are positive for MRSA.

[c] Prevalence ratios adjusted for sex and age.

[d] Tasks performed in the week before time 0.

[e] The task healthcare and control includes the administration of antibiotics.

[t] Nonsignificant trend (p-value 0.05-0.10). [*] p-value 0.01-0.05. [**] p-value 0.0001-0.01. [***] p-value <0.0001.

DISCUSSION

This study shows that the level of exposure to veal calves, administration of antimicrobials to animals and the presence of potential animal reservoirs (pets, free-ranging farm cats and sheep) are risk factors for persistent MRSA carriage in farmers and household members. MRSA persistent carriers seem to be characterised by higher MRSA CFU counts, presence of MSRA in throat and absence of MSSA. Carriage patterns of *S. aureus* presented are similar to those described by Wertheim and co-workers (2005) in which they found percentages of 20%, 30% and 50% for persistent, intermittent and non-carriers respectively among healthy individuals.

Persistent MRSA carriage is associated with higher MRSA CFU counts (PR=1.7 per 10 CFU increase). MRSA in throat swabs at day 0 was also significantly associated with persistent carriage (PR=12.2). This is suggestive of true colonisation in persistent MRSA carriers. A recent study shows that ST398 is capable of adequately competing for a niche with a human strain and survives in the human nose for longer periods (Slingerland et al., 2012).

Direct association between administration of antibiotics and short-term persistent carriage in humans was shown (PR=3.2). Where these drugs have been administered they can be present in manure, on surfaces of animal houses and in dust as a potential risk source (Hamscher et al., 2003). Aspiration of dust containing antibiotics either from a contaminated environment or directly from a powder formulation might have occurred.

In recent years, several reports have suggested a potential role for pet animals, specifically cats and dogs, in MRSA transmission and relapse of human MRSA infections. This transmission seems to be of anthropozoonotic origin (Bramble et al., 2011). It is remarkable that having a pet in the household and the presence of free-ranging farm cats and sheep were strongly associated with MRSA carriage in veal farmers and household members. These animals might represent an intermittent source of LA-MRSA that might contribute to LA-MRSA persistent carriage in humans, but this requires further study.

Other farm characteristics and hygiene practices were also associated with persistent MRSA carriage, although not significantly. Larger studies are required to confirm these findings. A possible limitation of the study is the self-sampling of nose and throat by individuals which might be lacking in accuracy of MRSA detection. However, a recent pilot study has shown high degree of agreement between self-samples and investigator samples (93% agreement, kappa 0.85 for nasal swabs and 83% agreement, kappa 0.60 for throat swabs) (van Cleef et al., 2012).

In conclusion, longitudinal associations found in this study give improved insights into possible intervention strategies to control MRSA on the farm. The increased risk for persistent MRSA carriage when there is a preceding administration of antimicrobials could have an impact in the discussion on the use of non-therapeutic treatments in livestock. Additionally the associations found here with pets, free-ranging farm cats and multispecies farming ask for detailed molecular-epidemiological analysis of MRSA specimens on the farm in various animal species and humans to identify reservoirs and transmission routes for LA-MRSA.

ACKNOWLEDGEMENTS

This study was funded by The Netherlands Organization for Health Research and Development (ZonMw) (grant number 125020009) and the Product Board for Livestock and Meat.

REFERENCES

Bramble, M., Morris, D., Tolomeo, P. and Lautenbach, E. (2011). Potential role of pet animals in household transmission of methicillin-resistant Staphylococcus aureus: a narrative review. Vector Borne Zoonotic Dis. 11, 617-620

Fitzgerald, J.R. (2012). Livestock-associated Staphylococcus aureus: origin, evolution and public health threat. Trends Microbiol. 20, 192-198

Graveland, H., Duim, B., van Duijkeren, E., Heederik, D. and Wagenaar, J.A. (2011a). Livestock-associated methicillin-resistant Staphylococcus aureus in animals and humans. Int. J. Med. Microbiol. 301, 630-634

Graveland, H., Wagenaar, J.A., Bergs, K., Heesterbeek, H. and Heederik, D. (2011b). Persistence of livestock associated MRSA CC398 in humans is dependent on intensity of animal contact. PloS One 6, e16830

Graveland, H., Wagenaar, J.A., Heesterbeek, H., Mevius, D., van Duijkeren, E. and Heederik, D. (2010). Methicillin resistant Staphylococcus aureus ST398 in veal calf farming: human MRSA carriage related with animal antimicrobial usage and farm hygiene. PloS One 5, e10990

Hamscher, G., Pawelzick, H.T., Sczesny, S., Nau, H. and Hartung, J. (2003). Antibiotics in dust originating from a pig-fattening farm: a new source of health hazard for farmers? Environ. Health Perspect. 111, 1590-1594

Rasigade, J.P., Laurent, F., Hubert, P., Vandenesch, F. and Etienne, J. (2010). Lethal necrotizing pneumonia caused by an ST398 Staphylococcus aureus strain. Emerg. Infect. Dis. 16, 1330

Slingerland, B.C., Tavakol, M., McCarthy, A.J., Lindsay, J.A., Snijders, S.V., Wagenaar, J.A., van Belkum, A., Vos, M.C., Verbrugh, H.A. and van Wamel, W.J. (2012). Survival of Staphylococcus aureus ST398 in the Human Nose after Artificial Inoculation. PloS One 7, e48896

van Cleef, B.A., Monnet, D.L., Voss, A., Krziwanek, K., Allerberger, F., Struelens, M., Zemlickova, H., Skov, R.L., Vuopio-Varkila, J., Cuny, C., Friedrich, A.W., Spiliopoulou, I., Paszti, J., Hardardottir, H., Rossney, A., Pan, A., Pantosti, A., Borg, M., Grundmann, H., Mueller-Premru, M., Olsson-Liljequist, B., Widmer, A., Harbarth, S., Schweiger, A., Unal, S. and Kluytmans, J.A. (2011). Livestock-associated methicillin-resistant Staphylococcus aureus in humans, Europe. Emerg. Infect. Dis. 17, 502-505

van Cleef, B.A., van Rijen, M., Ferket, M. and Kluytmans, J.A. (2012). Self-sampling is appropriate for detection of Staphylococcus aureus: a validation study. Antimicrob. Resist. Infect. Control 1, 34

van der Mee-Marquet, N., Francois, P., Domelier-Valentin, A.S., Coulomb, F., Decreux, C., Hombrock-Allet, C., Lehiani, O., Neveu, C., Ratovohery, D., Schrenzel, J., Quentin, R. and Bloodstream Infection Study Group of Reseau des Hygienistes du Centre (RHC)

(2011). Emergence of unusual bloodstream infections associated with pig-borne-like Staphylococcus aureus ST398 in France. Clin. Infect. Dis. 52, 152-153

van Duijkeren, E., Box, A.T., Heck, M.E., Wannet, W.J. and Fluit, A.C. (2004). Methicillin-resistant staphylococci isolated from animals. Vet. Microbiol. 103, 91-97

Wertheim, H.F., Melles, D.C., Vos, M.C., van Leeuwen, W., van Belkum, A., Verbrugh, H.A. and Nouwen, J.L. (2005). The role of nasal carriage in Staphylococcus aureus infections. Lancet Infect. Dis. 5, 751-762

SYNDROMIC SURVEILLANCE

USING BOVINE MEAT INSPECTION DATA FOR SYNDROMIC SURVEILLANCE:

INNOVATIVE STATISTICAL APPROACH FOR DEFINING SYNDROMES

C. DUPUY[*], E. MORIGNAT, P. HENDRIKX, C. DUCROT, X. MAUGEY, JL. VINARD, D. CALAVAS AND E. GAY

SUMMARY

Defining syndromes when performing syndromic surveillance is challenging. Definitions are commonly based on expert opinion consensus, not always easy to reach. Syndrome definition based on meat inspection data is particularly challenging due to the type of data collected regarding zootechnical information (breed, age, sex) and condemnation (portions and motives). Furthermore, each animal can have more than one condemnation motive. This study presents a statistical approach for syndrome definition based on both Multiple Factorial Analysis and clustering methods to define groups of condemnation motives, portions and zootechnical characteristics. Data from ten French bovine slaughterhouses were used, involving 1,937,917 cattle. Twelve clusters considered as stable according to year of slaughter and slaughterhouse were identified that deal with animal or public health surveillance, but also animal welfare, quality of the slaughtering process, management issues and evaluation of economical impact on farmers.

INTRODUCTION

Slaughterhouses, a focal point of farm animals, are a potential source of data not otherwise available. Both zootechnical data (age, breed, sex) and condemnation data (condemnation motives and portions) can be collected. At this time, meat inspection data are under-used for purposes other than food safety. The recent development of slaughterhouse databases in several countries makes these data accessible for cattle health monitoring (Alton et al., 2010; Ruoho et al., 2010; Dupuy et al., 2012b).

However, as the main objective of meat inspection is food safety and not animal health surveillance, the aetiological diagnosis is often not reached. Meat inspection data have a variable specificity from non specific lesions (e.g. abscess) to very specific ones (e.g. liver fluke).

---

[*] Céline Dupuy, Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail, 31 avenue Tony Garnier, Lyon, France. Email:celine.dupuy@anses.fr

Meat inspection and epidemiological surveillance

According to the type of meat inspection data used, different types of epidemiological surveillance could be performed (Table 1). Targeted surveillance (i.e. targeted objective) is focused on the surveillance of a pre-defined disease or group of diseases whereas non-targeted surveillance (i.e. non-targeted objective) aims at detecting unknown or emergent diseases. Syndromic surveillance can be either targeted or non-targeted according to the nature of the indicator monitored.

Table 1. Types of epidemiological surveillance according to the objective and examples of syndrome definition based on meat inspection data

Targeted objective	Non-targeted objective
Syndromic surveillance *e.g. lung lesions for respiratory diseases*	Syndromic surveillance *e.g. proportion of condemned cattle*
Traditional surveillance *e.g. muscular cysticercosis, fascioliasis*	

Impact of the type of epidemiological surveillance on syndrome definition method

The methodology used to define a “case” depends on the type of surveillance. For traditional surveillance, the case definition method is based on existing referential or expert definition, then relevant data are collected to meet the objective. For syndromic surveillance, available data are usually collected for another purpose and the epidemiologist usually has no power to modify the way data are collected, so using these data can be challenging. No referential of syndrome definition based on meat inspection data exists. Referring to complexity of the data (i.e. large diversity of motives and locations of condemnations), the definition through an expert process is difficult. The objective of this study was to implement a statistical approach to define groups of animals according to their zootechnical characteristics and lesions observed during the slaughtering process that could be used to define relevant syndromes to monitor.

MATERIALS AND METHODS

Data from ten French bovine slaughterhouses collected in real-time during the slaughtering process were available through the *Nergal-Abattoir* project implemented by the French ministry of Agriculture. Among the 1,937,917 cattle slaughtered in these slaughterhouses from July 2005 to December 2010, 381,186 (19.7%) had at least one part of the carcass condemned and were included in this study. For each bovine, the database contained: identification number, dates of birth and slaughter, last farm location, sex, breed, signs observed during ante-mortem inspection, condemnation motives and locations. All information registered during the slaughtering process and used to create variables for analyses are presented in Table 2. The ages of cattle were classified into six categories: 0-8 months, 8-24 months old, 2-3.5 years old, 3.5-5 years old, 5-10 years old and ≥10 years old. Some modalities of variables were grouped to avoid low numbers in categories that could create instability in data analyses. The 264 different condemnation motives of the *Nergal-Abattoir* system were merged into 57 categories of condemnation motives according to their biological similarities to be able to compare data among slaughterhouses when the levels of detail of the motive were different (e.g. abscess, multiple abscess, local abscess were merged into abscess). The condemnation portions were merged into 12 modalities (whole carcass, part of carcass, skin, head, udder, heart, lungs, liver, kidneys, viscera, tongue and thymus).

Table 2. Steps of the slaughtering process, data collected and variables created for analysis

Slaughtering step	Data collected	Variables
Identification check : Registration of zootechnical information	Animal ID/Farm ID	Farm location (Region)
	Sex (Male/Female) Castrated (Yes/no)	Sex: Female/Castrated male/Non-castrated male
	Breed (67 modalities)	Production type (Dairy/Beef/Mixed)
	Date of birth, Date of slaughter	Age category, Year and Month of slaughter
Ante-mortem inspection (AMI)	Signs observed (42 modalities)	AMI abnormality : Yes/No
Post-mortem inspection (PMI)	Condemnation motives: one for each location (264 modalities)	Condemnation motives (57 modalities)
	Condemnation portion (916 modalities)	Condemnation portions (12 modalities)

Multiple Factor Analysis (MFA) associated with clustering methods was used to establish groups of condemned cattle. Calculations were performed with R software (R Development Core Team, 2010), “FactomineR” and “cluster” packages. The steps of data analysis are presented in Fig.1.

Active variables that contributed to the construction of factorial axes of MFA were both zootechnical characteristics (sex, age category, production type) of condemned cattle and condemnation data (motives and portions). Year of slaughter, month of slaughter, farm location, presence of clinical signs during ante-mortem inspection, abattoir identification number, condemnation motives and portions with a proportion lower than 1% were used as illustrative variables i.e. variables that did not contribute to the construction of factorial axes but helped interpreting results of MFA. Clustering of bovines was performed from their principal coordinates by computing the classical Euclidean distance between these coordinates (Bécue-Bertaut & Pagès, 2008). Considering the large amount of data a previous K-means was performed before Hierarchical Ascendant Clustering (HAC). The generalised Ward’s criterion was used. The partition was determined considering the hierarchical tree and according to the biological meaning of the clusters (Lebart et al., 2006). The HAC was consolidated by a K-means performed on the centers of the HAC clusters.

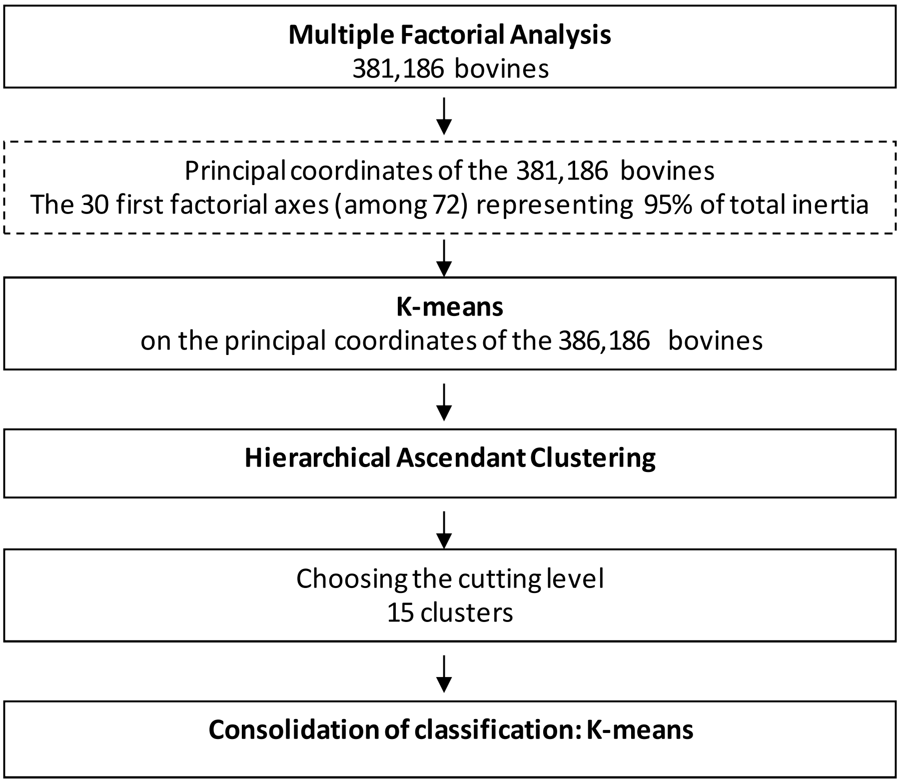


Fig.1 Steps of statistical data analysis

Stability of defined clusters was evaluated on 15 subsets of data according to year of slaughter (2006 to 2010: 5 subsets) and slaughterhouse (10 subsets) through the same process of analyses, i.e. MFA, K-means and HAC consolidated by K-means. Stability was defined as the proportion of subset analyses in which such a cluster was identified. Clusters with stability higher than 50% were considered as stable.

Description and interpretation of the clusters were based on both active and illustrative variables using V-Test values (Lebart et al., 2006). V-test of each modality of variables refers to the comparison between the proportion of cattle concerned by this modality in the cluster and the proportion of cattle with the same modality in the whole dataset formulated by a number of standard deviations of Gaussian law. Thus, the higher the difference between these two proportions is, the higher the absolute value of the V-test.

RESULTS

Descriptive analysis

According to the slaughterhouse, the mean number of cattle slaughtered each day varied from 122 to 543. The proportion of bovines having at least one portion of the carcass condemned varied from 9.6% to 35.9% according to slaughterhouses. Among the 381,186 cattle included in the study, (i.e. with condemnation), 70.0% were female, 25.5% non-castrated male and 4.5% castrated male. Beef cattle represented 44.3% of condemned cattle, dairy cattle 35.3% and mixed cattle 20.5%. The highest age category of condemned cattle was the 5- to 10-year-old age category (37.4%). The description of the 12 listed condemnation portions showed that liver condemnation concerned 67.5% of the cattle with at least one portion of the carcass condemned. The condemnation of the whole carcass concerned 3.4% of condemned cattle. The most frequent condemnation motives were abscess (18.8%) and liver fluke (15.1%).

Multiple factor analysis and clustering

Results of the different steps of the data analysis suggested four possible partitions in 6, 9, 15 or 16 clusters based on the height of the HAC dendrogram. The 15 clusters partition was selected based on the biological significance of the clusters.

The clusters 1, 2, 4, 5, 7, 8, 9, 10, 12, 13, 14 and 15 gathered between 1.3% and 2.4% of all condemned cattle. Clusters 6 and 11 were bigger, with respectively 4.1% and 6.0% of the condemned cattle. Cluster 3 gathered almost two thirds (65.7%) of the condemned cattle. Among the 15 clusters, clusters 5, 6 and 10 were considered as not stable and are not presented in this paper. The main variables that characterised the other 12 clusters are presented in Table 3.

Table 3. Main characteristics of the 12 stable clusters

Cluster	Main characteristic variables	Cla/Mod[a]	Mod/Cla[b]	Global[c]	|v.test|
1 (n=9,268)	Fecal contamination	82.4	100	3.0	+∞
	Lungs	10.9	81.3	18.1	+∞
	Heart	14.8	61.5	10.1	+∞
2 (n=7,215)	Bronchopneumonia	15.4	28.7	3.5	+∞
	Pericarditis	87.1	100	2.2	+∞
	Lungs	4.8	46.0	18.1	+∞
	Heart	18.7	99.7	10.1	+∞
	Dairy cattle	2.8	51.3	35.3	28.2
3 (n=250,256)	Macular telangiectasia	87.7	11.1	8.3	+∞
	Liver fluke	89.0	25.6	18.9	+∞
	Sclerosis	81.0	12.5	10.1	+∞
	Abscess	74.5	26.7	23.6	+∞
	Liver	76.9	79.1	67.5	+∞
4 (n=8,736)	Peritonitis	63.5	100	3.6	+∞
	Abscess	4.2	43.0	23.6	+∞
	Part of the carcass	9.7	54.2	12.8	+∞
7 (n=5,120)	Local muscular cysticercosis	99.8	100	1.4	+∞
	Head	12.9	75.4	7.9	+∞
	Heart	4.4	32.8	10.1	+∞
	(2 -3.5 years old)	3.0	29.9	13.3	31.3
	Castrated male	4.7	15.5	4.5	30.5
8 (n=9,095)	Other deteriorations[d]	86.8	100	2.8	+∞
9 (n=8,993)	Steatosis	83.9	100	2.8	+∞
	Kidneys	4.6	37.8	19.2	+∞
	Liver	3.1	87.9	67.5	+∞
	Dairy cattle	4.0	60.2	35.3	+∞
	(5-10 years old)	3.3	52.5	37.4	29.5
11 (n=22,690)	Ante-mortem abnormality	27.8	18.3	3.9	+∞
	Arthritis	86.1	43.1	3.0	+∞
	Infiltration	83.9	49.0	3.5	+∞
	Part of the carcass	46.1	99.2	12.8	+∞
12 (n=7,383)	Myopathy	94.9	100	2.0	+∞
	Part of the carcass	15.1	99.7	12.8	+∞
	Ante-mortem abnormality	5.3	10.7	3.9	25.3

13 (n=8,778)	Bronchopneumonia	65.3	100	3.5	+∞
	Lungs	12.7	99.8	18.1	+∞
	(8- 24 months old)	5.9	52.0	20.3	+∞
	Non-castrated male	5.0	55.3	25.5	+∞
14 (n=7,944)	Emphysema	94.2	100	2.2	+∞
	Lungs	11.5	100	18.1	+∞
	Dairy cattle	4.2	71.4	35.3	+∞
	(5-10 years old)	3.1	55.1	37.4	32.3
15 (n=5,452)	Meat with abnormal maturation	98.1	100	1.5	+∞
	Ante-mortem abnormality	11.1	30.2	3.9	+∞
	Whole carcass	41.6	99.8	3.4	+∞

[a] Proportion of cattle having this modality that were in the cluster e.g. 15.4% of cattle that had a lesion of bronchopneumonia in the whole population were in the cluster 2.

[b] Proportion of cattle in the cluster having the modality e.g. 28.7% of the cattle in the cluster 2 had a lesion of bronchopneumonia

[c] Proportion of cattle in the global population having the modality e.g. 3.5% of the whole population of cattle with condemnation had a lesion of bronchopneumonia.

[d] Including haemorrhagic lesions of lungs or muscles linked to problems in the slaughtering process; superficial and deep putrefaction.

DISCUSSION

The objective of this study was to determine a typology of bovines with condemnation in order to define syndromes based on a statistical approach.

Clusters interpretation: typology of condemned cattle

The interpretation of each of the 12 clusters showed that meat inspection data can be used beyond food safety. The clusters describing various types of condemnation referred to five different fields of interpretation in relationship with farming and abattoir activities: animal health, animal welfare, breeding practice, public health and quality of the slaughtering process (Table 4). The three types of epidemiological surveillance in Table 1 could be addressed by monitoring these syndromes: traditional surveillance with cluster 7 (cysticercosis), targeted syndromic surveillance with for instance cluster 4 (chronic peritonitis lesions), non-targeted syndromic surveillance by monitoring the percentage of cattle in each cluster to detect changes in trends.

Limits of this typology

The typology of lesions presented in this study was based on French meat inspection data and was thus linked to the animal health and welfare situation in France. It could be interesting to build a similar typology based on meat inspection data from other countries to (i) highlight common clusters (i.e. groups of lesions) for which exchanges of information and knowledge could be relevant and (ii) try to understand why adjacent countries do not always share the same clusters.

As disease presence and patterns can change over years within a country, it could be useful to perform the description of groups of lesions based on the statistical method presented again when the health status of the country seems to have changed or after a defined period of time. This change would be detected by an important modification in trends of the proportion of cattle in each cluster.

Table 4. Interpretation and field of interest of stable statistical clusters identified

Cluster	Interpretation	Field
1	Fecal contamination of heart or lungs linked with failure during evisceration step	Quality of slaughtering process
8	Deteriorations linked to failure in the slaughtering process	
2	Pericarditis and bronchopneumonia which could be linked to traumatic reticulo-pericarditis	Breeding practice such as feeding
9	Fatty liver syndrome	
14	Lung emphysema : farmer’s lung	
3	Chronic liver lesions	Breeding practice
4	Chronic peritonitis lesions which could be linked to previous infection	
11	Arthritis	Animal health, Animal welfare
12	Myopathy (muscle crush syndrome)	Animal welfare, Breeding practice
13	Bronchopneumonia	Animal health, Breeding practice
15	Dark Firm Dry meat	Animal welfare (transport management)
7	Cysticercosis	Public health (zoonotic disease)

Syndrome definition methods according to the type of syndromic surveillance

As stated in the introduction, the syndrome definition method is different according to the type of syndromic surveillance (Table 5). For targeted syndromic surveillance (i.e. the objective of the system is the detection of a known disease or group of diseases), three approaches could be used: (1) definition based on an existing referential which, to our knowledge, does not exist for meat inspection data; (2) definition based on expert opinion consensus; and (3) definition based on both expert opinion and statistical typology of existing groups of lesions performed on historical data. Indeed reaching a consensus among experts based on a wide list of lesions associated or not with zootechnical criteria is challenging. It seems relevant to start with a statistical analysis of groups of lesions and zootechnical characteristics as support for further discussion among experts.

For non-targeted syndromic surveillance (i.e. the objective of the system is the detection of unknown diseases), generic indicators need to be defined. Two approaches have been identified. First, based on the hypothesis that each disease has a dedicated organ tropism, more or less specific, and evolution stages of lesions, more or less acute, it seems relevant to assume that monitoring the proportion of condemned cattle according to the extent/location of condemnation (whole carcass, part of the carcass, heart, liver…), whatever the nature of the lesion, could allow the detection of an emerging disease. Second, based on the fact that each disease is characterised by group(s) of lesions and that it affects a susceptible bovine population characterised by zootechnical criteria, it seems relevant to assume that most cattle affected with an emerging disease will have similar lesions and zootechnical characteristics. Using the statistical syndrome definition, each condemned bovine added in the dataset used to define clusters could be attributed to a cluster by determining its MFA-derived representation and the cluster it belongs to. It is thus probable that most of these cattle will be attributed to the same cluster, making its proportion abnormally increasing.

Table 5. Advantages and limits of the different syndrome definition methods

Type of syndromic surveillance	Syndrome definition method	Advantages	Limits
Targeted	Based on existing referential	Easy to use Easier for comparability among systems Cheap	Need of an existing referential Accessibility of the data as defined in the referential
	Based on expert opinion consensus	Classical and known approach	Difficulty to reach consensus Depends on the experts recruited Cost for expert group Accessibility of the data as defined by experts
	Statistical definition associated with expert biological interpretation	Based on both historical data and expert opinion Help for reaching consensus Could identify syndromes not identified by experts Start point is available data which guarantee accessibility to data needed	Historical data needed Depends on the experts recruited Cost for expert group
Non-targeted	Statistical definition	Based on historical data/the reality Possible for complex and large sets of data	Historical data needed
	Expert definition: Generic indicator	Easy to define	Not able to take into account the complexity of the data

Added value and limits of meat inspection data for bovine surveillance

Using meat inspection data for animal health surveillance is tempting. Indeed these data are not otherwise available which makes the detection of diseases not detectable by other types of surveillance (e.g. cysticercosis) possible. The slaughterhouse is the focal point of most cattle thus meat inspection data deals with nearly the entire bovine population. More and more countries are implementing databases to collect these data. This could allow the implementation of a nearly exhaustive surveillance system. As demonstrated in this study, meat inspection data deals not only with food safety but also with animal health and welfare, breeding practice and the slaughtering process, making these data potentially useful for several objectives and stakeholders (veterinary services, farmers, etc.).

Limits of meat inspection data could however be identified. As explained before, these data are collected for food safety purpose which has an impact on the specificity of these data if used for another purpose. It is also difficult to make the link between the lesions observed at the slaughterhouse and the location of the onset of the disease (last farm or not?) especially for chronic lesions. It could create important biases in spatial analyses if it were not taken into account. As with all systems collecting data through different persons, there could be differences according to the person performing the meat inspection. Likewise, differences could exist according to the slaughterhouse in terms of sensitivity and specificity of the inspection. An EFSA European project called “Contribution of meat inspection to animal health surveillance” was performed to quantify the effectiveness of meat inspection to detect a list of diseases and welfare conditions (Dupuy et al., 2012a). The probability of detection of a disease (i.e. probability of detection of a sign during ante-mortem inspection or lesion during post-mortem inspection) depends on a lot of factors that need to be taken into account when using meat inspection data such as the disease stage (acute or chronic), the training level of the inspector and the prevalence of the disease in this area (i.e. lesion more or less known by inspectors).

Finally, as demonstrated before, a huge amount of data are available from slaughterhouses with a lot of different lesions, condemnation portions and zootechnical information that make definition of relevant groups of lesions difficult. The statistical approach presented in this study is one of the answers to this issue.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the French Ministry of Agriculture for funding this study and for having provided access to the data. We also wish to thank the group of experts from the Animal Société Aliment Association (ASA) (http://www.asa-spv.asso.fr/) and Stéphane Pierret, veterinary inspector in a French slaughterhouse who helped with typology interpretation.

REFERENCES

Alton, G., Pearl, D., Bateman, K., McNab, W. and Berke, O. (2010). Factors associated with whole carcass condemnation rates in provincially-inspected abattoirs in Ontario 2001-2007: implications for food animal syndromic surveillance. BMC Vet. Res. 6, 42

Bécue-Bertaut, M. and Pagès, J. (2008). Multiple factor analysis and clustering of a mixture of quantitative, categorical and frequency data. Comput. Stat. Data Anal. 52, 3255-3268

Dupuy, C., Hendrikx, P., Hardstaff, J. and Lindberg, A. (2012a). Contribution of meat inspection to animal health surveillance in Bovine animals. In: EFSA (Ed.) European Food Safety Authority, p. 53

Dupuy, C., Morignat, E., Gay, E. and Calavas, D. (2012b). Risk factors for condemnation in cattle slaughtered in a French abattoir from 2006 to 2009. In : XXVII World Buiatrics Congress, Lisbon, Portugal

Lebart, L., Morineau, A. and Piron, M. (2006). Statistique exploratoire multidimensionnelle, visualisation et inférence en fouilles de données. DUNOD Paris

R Development Core Team (2010). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org

Ruoho, O., Kortesniemi, P. and Halkosaari, P. (2010). Transferring data from farm to slaughterhouse "on-line" via centralized register. XXVI World Buiatrics Congress. Santiago, Chili

SYNDROMIC SURVEILLANCE IN VETERINARY MEDICINE USING LABORATORY SUBMISSION DATA – LESSONS LEARNED FROM TWO SYSTEMS

F. C. DÓREA[*], A. LINDBERG, B. J. MCEWEN, C.W. REVIE AND J. SANCHEZ

SUMMARY

Syndromic surveillance can be characterised as a process involving the continuous analysis of health data to provide immediate feedback. The data sources scanned should therefore represent high population coverage, be acquired continuously, in an automated manner, and be available in digital format. Data from diagnostic test requests often meet these requirements. Building on the experiences from developing syndromic surveillance in two institutions, the steps from data extraction to eventual aberration detection are described in this paper, and can be summarised as: classification of records into syndromes; retrospective evaluation of data to create baseline profiles following the removal of excessive noise and aberrations, and the identification of temporal effects; prospective evaluation of detection algorithms; and finally real-time monitoring and implementation. All steps described were implemented using open source software, and could be readily reproduced in other institutions.

INTRODUCTION

Syndromic surveillance is a tool for continuous, automated extraction of surveillance information from health data sources, as timely as the rate of gathering data in electronic format. Timely collection and computerisation were some of the reasons highlighted by Ma et al. (2005) for the use of laboratory order data in syndromic surveillance, along with their high population coverage. The scarcity of centralised electronic collections of clinical data in veterinary medicine further motivates the use of laboratory data in animal health surveillance (Dórea et al., 2011).

After a decade of syndromic surveillance development in public health, the focus today is on developing holistic biosurveillance systems which work in parallel – rather than as a substitute – and help inform and complement other types of surveillance (Buckeridge, 2011; Fricker Jr, 2011). Authors, who in the last decade researched the potential of data sources and the use of different aberration detection algorithms, now highlight the need to continue research into system design and implementation. They also point out the need to compare methods developed under different conditions and to document their performance, fomenting sharing experiences in order to enhance syndromic surveillance utility (Fricker Jr, 2011).

---

[*] Fernanda C. Dórea. Swedish Zoonosis Centre, Department of Disease Control and Epidemiology, National Veterinary Institute (SVA). SE-751 89 Uppsala, Sweden.
Email: fernanda.dorea@sva.se.

The objective of this paper is to present a structured approach for designing and implementing syndromic surveillance based on laboratory test requests to two animal health laboratories with high population coverage. These laboratories were the Animal Health Laboratory (AHL) in the province of Ontario, Canada, and the laboratory at the Swedish National Veterinary Institute (SVA). The latter system is still in its development phase, but comparisons between the two data sources will be made as well as an assessment of the portability of the system and the challenges of working with each data source.

MATERIALS AND METHODS

An overview of the process of developing a syndromic surveillance system from animal health data (with a focus on structured tabular data, such as laboratory data) is presented schematically in Fig. 1. Details of each step are described below.

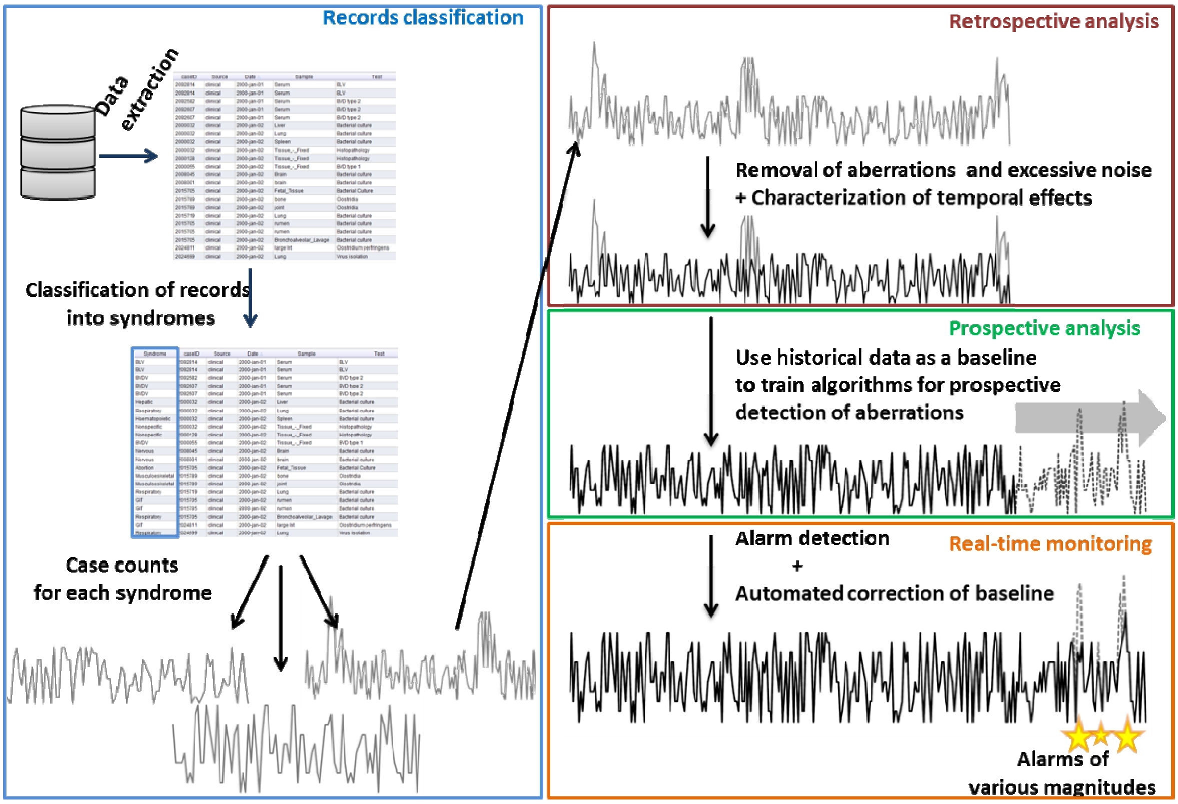


Fig. 1 Overview of the process of developing a syndromic surveillance system from animal health data

Data sources

Two syndromic surveillance systems were developed on the initiative of the data owners, with all phases of data access, analysis and output generation developed inside the institution. The AHL, part of the University of Guelph, is a full-service veterinary diagnostic laboratory that serves livestock, poultry and companion animal veterinarians in the province of Ontario, Canada. The SVA is an agency under the Ministry of Rural Affairs and is the largest and only official laboratory in veterinary medicine in the country. Test requests in both laboratories are digitalised daily, data and operations are managed by a laboratory information management system (LIMS). The format of this information is therefore very structured. Most data fields which can be used for surveillance purposes – such as samples submitted and diagnostic tests performed – have a pre-set list of options and little free-text is entered. Only cattle data were used at the AHL, whilst data from all species are being included at the SVA.

The institutes' diagnosticians were involved in defining the data fields relevant for the extraction of surveillance information. Only data available at the time of submission were considered. Fields with information concerning animal owners and veterinarians were not included in the data extracted for reasons of confidentiality. Only data submitted to the laboratories by field veterinarians requesting diagnostic tests were considered relevant for the purpose of syndromic surveillance.

Automated identification of cases and syndromes

Once data sources have been identified, the next step is to define how cases can be extracted. Syndrome definitions and ways of implementing automated classification of the computerised records must be researched. Rather than defining the syndromes a priori, an exhaustive list of syndromic groups was created, so that every case could be classified to at least one syndrome. Experts from the laboratories were asked to contribute to this process.

Retrospective analysis

Evaluation of temporal effects and time series modelling: Once the syndromic groups were defined, statistical properties of each syndromic time series were evaluated individually in both of the systems developed. All statistical analyses were performed in the R environment (http://www.r-project.org/) (R Core Team, 2012). Health data are typically subjected to temporal effects such as day-of-week, season, global trends, as well as influence of various random factors not related to disease burden. The challenges of continuously monitoring health data and dealing with such effects have been thoroughly discussed in the literature (e.g. Yahav & Shmueli, 2007; Lotze et al., 2008; Shmueli & Burkom, 2010; Dórea et al., 2012).

Cleaning noise and historical outbreaks: To detect temporal aberrations in real-time, observed data can be compared to historical data known to be free of outbreaks (data-driven methods) or to a model reproducing the expected behaviour (model-driven methods) (Yahav & Shmueli, 2007). In both cases, a baseline representing the expected behaviour in the absence of outbreaks is necessary. Tsui et al. (2001) proposed a method for cleaning historical data based on the removal of observations above the upper limit of a 95% confidence interval. The goal in both systems developed was to generalise the method so that it could be applied to time series with different statistical properties.

Prospective analysis

Algorithms for detection of temporal aberrations: For comprehensive reviews of the use of temporal detection algorithms in biosurveillance refer to Buckeridge et al., 2005; Burkom & Murphy, 2007; Yahav & Shmueli, 2007; Höhle et al., 2009; Shmueli & Burkom, 2010; and references therein. Classical quality control methods have been used extensively in biosurveillance (Woodall, 2006). Control charts are based on the assumption that observations are independently drawn from pre-specified parametric distributions, and therefore pre-processing of health data is necessary in order to remove the temporal effects identified (Lotze et al., 2008). An alternative is to make use of data-driven statistical methods, such as Holt-Winters exponential smoothing, which can efficiently account for temporal effects (Elbert & Burkom, 2009). Three control charts were explored: Cumulative sums (CUSUM), Exponentially Weighted Moving Averages (EWMA) and Shewhart control charts (Woodall, 2006). A data-driven method was also explored, using Holt-Winters exponential smoothing (Elbert & Burkom, 2009). Two years of historical data were used to train the algorithms used in these systems, while the remaining available data (at least 2 years) were used to simulate forward monitoring, in order to qualitatively evaluate the performance of different algorithms in the detection of temporal aberrations as well as in the generation of false alarms.

Define detection limits and evaluate algorithm performance: The main parameter affecting detection of temporal aberrations is the detection limit set for the algorithm. For control charts this detection limit is based on standard deviations, while for the Holt-Winters exponential smoothing algorithm it is based on the confidence interval of a model prediction for the current time point. In order to construct standard receiver operating characteristic (ROC) curves, and define detection thresholds based on the desired balance between sensitivity and specificity, the various algorithms were tested using simulated data with controlled injections of outbreaks of various types and durations.

Real-time monitoring

When implemented in real-time, detection algorithms work in a “roll-forward” manner (Fienberg & Shmueli, 2005). It is recommended that a “guard-band” gap is left between the baseline used to estimate expected behaviour and the current time point, in order to avoid the situation in which slowly rising case, but not sufficient to be detected as an outbreak, contaminate the baseline (Burkom, 2003). The use of roll-forward algorithms was simulated to a start date at least one-year prior to the time of implementation, in order to evaluate system outputs and to tune algorithm parameters.

Implementation

Only freely available software was used. The software RapidMiner 5.0 (Copyright 2001-2010 by Rapid-I and contributors) was used to establish a connection to each institution’s database remotely, apply the SQL query to extract data, and execute the text mining and classification process. R codes (within the RapidMiner interface) were used to perform all statistical analysis, detect alarms, and generate outputs.

RESULTS

Data source

Data from AHL contained a specific data field to inform when sampling was part of a monitoring program (active sampling), submitted by a veterinarian for diagnostic (passive submissions), or part of research. In contrast, the main challenge in extracting data at SVA was that of differentiating passive from active sampling and research. An attempt was made to differentiate passive from active sampling based on the party billed. Samples paid from government sources were considered to be part of mandatory programmes. Some governmental funding can however be available as an incentive for submission of samples from sick animals. As much as possible keywords that would ensure the observation of a clinical problem in the field – such as keywords related to abortion (abortion, aborted, fetus, placenta, etc) or words such as "suspicion" or "notification" – were used to find and label samples that should be considered a result of passive submission, regardless of the payer. Missing information on animal species also hindered data extraction. Once again keywords search was used in order to look for information among all other fields available at submission. For instance, words like "bovine" and "swine" were searched for so that the species could be determined when the diagnostic test was such as "bovine leukemia" or "classical swine fever".

Automated identification of cases and syndromes

Individual tests are reported as unique data entries. In the AHL database a single submission code is assigned to samples submitted from the same herd in a given day, but the number of animals sampled is not recorded. For this database, therefore, a case was defined as one unique submission. At SVA it was possible to keep track of the number of animals and herds sampled per submission.

Due to the limited amount of clinical information, the final syndromic classification was based on organ systems targeted by the veterinarian for diagnosis, mainly: circulatory, hepatic and hematopoietic; gastrointestinal (GIT); integumentary; musculoskeletal; nervous; reproductive; respiratory; systemic; urinary and others. Specific groups to identify cases of abortion and mastitis were also created. Some groups were created in order to provide exhaustive classification of records, and accommodate submissions that could not be classified into a specific organ system, such as: antimicrobial resistance; clinical pathology tests; non-specific tests; whole carcass and toxicology tests.

In order to automate the classification process, a direct correspondence (mapping) was established between pathogen specific tests and syndromes (for instance rabies to nervous syndrome). In case of non-specific tests, such as "bacteriological investigation", or "histology", text mining methods were used to search the fields describing the specimens submitted. The highest classification performance was achieved when a set of rules linking keywords and syndromic groups was defined manually (rule-based classification) – 99.4% recall and 100% precision. Remaining errors were due to data quality issues, such as misspellings or use of ambiguous acronyms. This was found to be one of the most time consuming steps of developing the system.

When proceeding to develop a similar system at SVA, little could be borrowed from the previous system. Apart from the obvious language barrier, essential differences existed in the way tests and specimens were recorded. At SVA test codes are more often modified, and test names can be used more freely. For that reason, the first step of classification could not be fully performed by mapping, and rule-based classification was needed also for this step, requiring an extra dictionary of keywords and set of rules. In both systems syndromic classification was performed for each individual database entry (test request), and each case was counted in as many syndromes as the organ systems identified.

Retrospective analysis

Once all data were classified, time-series with the number of cases per day, for each syndromic group, were constructed and evaluated retrospectively.

Evaluation of temporal effects and series modelling: All time series in both systems showed strong day-of-week effect. The syndromic series at AHL showed weak seasonal effects and no global trends. All series with median counts greater than one case per day were successfully modelled using Poisson regression with month and day-of-week as predictors. Syndromes with counts lower than that were aggregated into weekly counts, and the resulting time series were not found to exhibit temporal patterns that could impact the performance of outbreak detection methods. The SVA data produced a much larger number of syndromes, and their time series showed a greater variety of temporal effects. Clear seasonal patterns were observed even in series with very low counts, especially for equine submissions.

A greater number of syndromes with very low daily counts were found in the SVA data, and in some cases even weekly aggregation would result in a large number of time points with zero observations. Methods were sought to deal with these sparse counts, and to generalise the methods for use with any syndromic series. Three options are being evaluated (Lucas, 1985): (i) using distributions appropriate for over-dispersed counts, such as negative binomial (Sparks et al., 2010); (ii) monitoring the time between two submissions (in days) (Radaelli, 1998); and (iii) monitoring the total number of submissions between each submission for a specific syndrome (Sego et al., 2008; Szarka & Woodall, 2011). The latter has the advantage of correcting for an overall increase in the number of submissions to the laboratory.

Removing noise and historical outbreaks: Using the AHL data it was possible to generalise the method proposed by Tsui et al. (2001) using Poisson regression, as previously documented in Dórea et al. (2012). In summary, a Poisson model was fit to the syndromic time series, values above the 95th percentile interval were substituted by the limit of that percentile, and the process was repeated iteratively until no observations fell outside the limit. For series grouped into weekly counts the process was based on the 95th percentile of observations in a 26-weeks moving window.

With the much larger number of time series generated in the SVA system, it was possible to generalise the process of data cleaning to three alternative methods: (i) find a regression model that satisfactorily fits the data, based on goodness-of-fit tests and analysis of residuals, and then apply the iterative steps of model fitting and removal of outliers described in the previous paragraph; (ii) remove temporal effects and subject residuals (assumed to be independently and identically distributed – i.i.d. – after temporal effects removal) to the non-parametric cleaning process using moving percentiles; (iii) apply the non-parametric procedure directly to the data after assuring that no reproducible temporal effects need to be removed.

Prospective analysis

Algorithms for detection of temporal aberrations: During development of the AHL system, it was found that Holt-Winters exponential smoothing, previously suggested as a pre-processing method to remove temporal effects (Yahav & Shmueli, 2007), performed better when used directly as an aberration detection algorithm. Alternatively, data were pre-processed by weekly differencing and then subjected to aberration detection using control charts. Exponentially Weighted Moving Averages (EWMA) and Shewhart control charts showed complementary performance in detecting outbreaks of different shapes, while cumulative sums (CUSUM) charts did not seem to improve performance once the first two control charts were already in use.

The development of the system at SVA has not yet reached the prospective phase, but preliminary evaluations indicate that only minor adjustments will be required to generalise the system built at AHL.

Define detection limits and evaluate algorithm performance: At AHL, evaluation using simulated data showed that no single algorithm performed better in all outbreak scenarios. In order to construct a system that takes advantage of the strengths of each detection algorithm, EWMA and Shewhart control charts, and the Holt-Winters exponential smoothing were combined into a scoring system. For each of these three algorithms five detection limits were applied, using scores from 1 (most sensitive) to 5 (most specific). In the event of an observed count being higher than any of these five detection limits, a score is registered corresponding to the highest detection limit exceeded. As a result, the system outcome for each syndrome, at each time point monitored, is a detection score between 0 and 15 (the maximum sum of the scores from all three algorithms), providing a progressive scale for the magnitude of any alarm. Analysts can tune the system by defining, individually for each of the syndromes, the score threshold that will trigger a system action (alarm). When applied to simulated data, the scoring system showed high sensitivity for detection of outbreaks, but also generated a high number of false alarms. However, when applied to real data, no more than one false alarm per month was documented.

Real-time monitoring

During the development of the system in Canada, prospective analysis was performed using data queried as a batch, but applying roll-forward algorithms. Only after the system had evaluated the full amount of historical data available actual querying of data in real-time was set up. The use of a guard-band was found to prevent baseline contamination only before the first detection of an outbreak, but not after. Once algorithms were implemented in a roll-forward routine, Holt-Winters exponential smoothing was set to correct any aberrations detected, continuously storing an outbreak-free baseline. The goal was to remove aberrations and excess noise, so the correction is applied anytime observed data is greater than the upper limit of a 97.5% confidence interval for the algorithm prediction, independently of the generation of an alarm in the combined scoring system. The observed count is then substituted by the limit of that confidence interval.

Implementation

The output in case of an alarm is an email to analysts involved in the project, containing summaries of the syndromic series as well as the detection score generated by each of the three algorithms. Every Friday the analysts receive emails with 26- and 52-week plots for all syndromes monitored, regardless of the generation of an alarm during the week.

DISCUSSION

The collection of specimens for diagnostics happens late in the disease continuum, and covers a small percentage of the population affected. For this reason, sensitivity and timely detection should be prioritised (Widdowson et al., 2003). In both systems developed, the data were monitored directly at the institution where they are collected and stored. A single, centralised, electronic source circumvents the need to gather data from scattered locations, as well as the need to develop standards that allow integration of these data from different sources. Issues of data transfer, sharing agreements and security are all also avoided if the data are monitored directly at the source institution.

The speed of extraction of surveillance information from health data does not depend solely on the rate of data collection, but also on the ability to translate data from the format they are collected to a meaningful surveillance output (Fienberg & Shmueli, 2005). The most time-consuming step of classification, in both systems, was the creation of a dictionary of keywords relevant to each classification task, and the definition of the relationship between these words, their co-occurrences and the target syndromic group. Once defined, however, these relationships were easily packed into a set of rules that achieved high classification performance.

Failures in automated classification of records could be tracked to two main problems: the secondary use of data stored for purposes other than surveillance – in the specific cases described here, for diagnostic purposes and management tasks; and data quality. Small artefacts can hinder the detection of cases by the automated classification algorithms, and result in a large impact in performance for the entire system (van den Wjingaard et al., 2011). As these systems are developed, feedback of these issues to the institution can help increase overall data quality.

One of the recommended parameters to evaluate in any surveillance system is its portability, that is, "how well the system could be duplicated in another setting" (Centers for Disease Control and Prevention (CDC), 2004). The steps to achieve syndromic classification showed no portability, but once all data were reduced to multiple time-series little difference existed between the steps necessary to handle data in the two different systems.

In both institutions training of analysts to read system outputs and take decisions regarding adjustment of detection thresholds was feasible. As the analysts become familiar with the system, the cycle of feedback between the analyst and the system will both refine the performance of the system and feed the institution with surveillance information which goes beyond early disease detection (Widdowson et al., 2003). If efforts are not taken inside the institutions to allocate time for this task, the full utility of the system will not be attained.

Both systems were developed as desktop applications at the institution that owns the data, scheduled to run daily. The structured development and implementation resulted in fast achievements for the projects. Some previous attempts to develop syndromic surveillance in veterinary medicine have failed to sustain engagement and become established systems (Dórea et al., 2011). The simplicity of the systems developed may facilitate their incorporation into institutions' routine activities, becoming a basic tool for diagnosticians and epidemiologists (Hurt-Mullen & Coberly, 2005).

ACKNOWLEDGEMENTS

The work at AHL was supported by the OMAFRA-UG Agreement through the Animal Health Strategic Investment fund (AHSI) managed by the Animal Health Laboratory of the University of Guelph. Development at SVA was funded by the Swedish Board of Agriculture.

REFERENCES

Buckeridge, D.L., Burkom, H., Campbell, M., Hogan, W.R. and Moore, A.W. (2005). Algorithms for rapid outbreak detection: a research synthesis. J. Biomed. Inf. 38, 99–113

Buckeridge, D.L. (2011). Comments on ‘some methodological issues in biosurveillance’. Stat. Med. 30, 420–422

Burkom, H.S. (2003). Development, adaptation, and assessment of alerting algorithms for biosurveillance. Johns Hopkins Apl. Tech. Dig. 24, 335–342

Burkom, H.S. and Murphy, S.P. (2007). Automated time series forecasting for biosurveillance. Stat. Med. 26, 4202–4218

Centers for Disease Control and Prevention (CDC) (2004). Framework for evaluating public health surveillance systems for early detection of outbreaks; recommendations from the CDC Working Group. MMWR, 53, 1-13

Dórea, F.C., Sanchez, J. and Revie, C.W. (2011). Veterinary syndromic surveillance: Current initiatives and potential for development. Prev. Vet. Med. 101, 1–17

Dórea, F.C., Revie, C.W., McEwen, B.J., McNab, W.B., Kelton, D. and Sanchez, J. (2012). Retrospective time series analysis of veterinary laboratory data preparing a historical baseline for cluster detection in syndromic surveillance. Prev. Vet. Med., in press, http://dx.doi.org/10.1016/j.prevetmed.2012.10.010

Elbert, Y. and Burkom, H.S. (2009). Development and evaluation of a data-adaptive alerting algorithm for univariate temporal biosurveillance data. Stat. Med. 28, 3226–3248.

Fienberg, S.E. and Shmueli, G. (2005). Statistical issues and challenges associated with rapid detection of bio-terrorist attacks. Stat. Med. 24, 513–530.

Fricker Jr, R.D. (2011). Some methodological issues in biosurveillance. Stat. Med. 30, 403–415

Höhle, M., Paul, M. and Held, L. (2009). Statistical approaches to the monitoring and surveillance of infectious diseases for veterinary public health. Prev. Vet. Med. 91, 2–10

Hurt-Mullen, K.J. and Coberly, J. (2005). Syndromic surveillance on the epidemiologist’s desktop: making sense of much data. MMWR 54 Suppl., 141–146

Lotze, T., Murphy, S. and Shmueli, G. (2008). Implementation and comparison of preprocessing methods for biosurveillance data. ADS 6, 1–20

Lucas, J.M. (1985). Counted data CUSUM’s. Technometrics. 27, 129–144

Ma, H., Rolka, H., Mandl, K., Buckeridge, D., Fleischauer, A. and Pavlin, J. (2005). Implementation of laboratory order data in biosense early event detection and situation awareness system. MMWR 54 Suppl., 27–30

R Core Team (2012). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0. http://¬www.R-project.org

Radaelli, G. (1998). Planning time-between-events Shewhart control charts. Total Qual. Manage 9, 133–140

Sego, L.H., Woodall, W.H. and Reynolds Jr, M.R. (2008). A comparison of surveillance methods for small incidence rates. Stat. Med. 27, 1225–1247

Shmueli, G. and Burkom, H. (2010). Statistical challenges facing early outbreak detection in biosurveillance. Technometrics 52, 39–51

Sparks, R., Keighley, T. and Muscatello, D. (2010). Early warning CUSUM plans for surveillance of negative binomial daily disease counts. J. Appl. Stat. 37, 1911–1929

Szarka, J.L. and Woodall, W.H. (2011). A review and perspective on surveillance of Bernoulli processes. Qual. Reliab. Eng. Int. 27, 735 – 752

Tsui, F.C., Wagner, M.M., Dato, V. and Chang, C.C. (2001). Value of ICD-9 coded chief complaints for detection of epidemics. AMIA Annu. Symp. Proc.1, 711–715

van den Wjingaard, C.C., van Pelt, W., Nagelkerke, N.J., Kretzschmar, M. and Koopmans, M. (2011). Evaluation of syndromic surveillance in the Netherlands: its added value and recommendations for implementation. Eurosurv. 16 (9)

Widdowson, M.-A., Bosman, A., van Straten, E., Tinga, M., Chaves, S., van Eerden, L. and van Pelt, W. (2003). Automated, laboratory-based system using the internet for disease outbreak detection, the Netherlands. Emerg. Infect. Dis. 9, 1046–1052

Woodall, W. (2006). Use of control charts in health-care and public-health surveillance. J. Qual. Tech. 38, 89–104

Yahav, I. and Shmueli, G. (2007). Algorithm combination for improved performance in biosurveillance systems. In: Gotham, I.; Komatsu, K.; Lynch, C.; Thurmond, M.; Madigan, D.; Lober, B.; Kvach, J. and Chen, H. (eds.) Biosurveillance, Springer, Berlin Heidelberg, 2007. pp. 91–102

SYNDROMIC SURVEILLANCE IN DAIRY CATTLE: DEVELOPMENT OF INDICATORS AND METHODS BASED ON REPRODUCTION DATA FOR EARLY DETECTION OF EMERGING DISEASES

A. MARCEAU, A. MADOUASSE, A. LEHEBEL, T. LESUFFLEUR, Y. VAN DER STEDE, G. VAN SCHAIK AND C. FOURICHON*

SUMMARY

Over the last decade, there has been an increase of disease emergence and in particular of vector borne diseases. In order to support specific surveillance or control programmes there is a need for syndromic surveillance. One way to achieve this is to utilise an extensively large amount of data which is produced frequently and automatically by veterinary health agencies, administration bodies and/or farmers' organisations. The objectives of this study were to develop and evaluate both a set of indicators based on reproduction data, and various statistical methods, to detect the emergence of a disease with known effects on reproduction in dairy cattle. The bluetongue outbreak in four districts in France in 2007 was used as a case model. Five reproductive indicators based on occurrence of return-to-service or gestation length were proposed and evaluated. Using the indicator with the highest elevations during the epidemic, five statistical models combined with a CUSUM (cumulative sum of differences between expected and observed data) method to trigger alarms were implemented and evaluated. Four indicators were able to flag the epidemic in all infected districts except one in which the cumulated yearly prevalence was very low. The indicator based on short gestation had the longest and the earliest elevations from prediction. The alarm system was effective in detecting the epidemic with all statistical models. Performance of the alarm system was good and varied with the threshold parameter value of the CUSUM method. A good compromise between a low number of false alarms (2-4 alarms over 4 years) and early timeliness (~3 weeks) was obtained for a parameter around 0.015.

INTRODUCTION

In response to increasing risks of disease emergences, syndromic surveillance can be a suitable approach to detect outbreaks of various adverse health events in an early phase. Such a system relies on real-time analysis of routinely collected data which are used as non-specific indicators of health status. As an example, fever and respiratory syndromes reported in US hospitals are used to detect the seasonal epidemic of influenza in human populations (Olson et al., 2007).

---

* Christine Fourichon, Oniris, UMR1300 Biologie, Epidémiologie et Analyse de Risque en santé animale, La Chantrerie, CS 40706, F-44307 Nantes, France.
Email : christine.fourichon@oniris-nantes.fr

To implement a syndromic surveillance system, the first challenge is to find relevant data. The data relevance is determined by their close association with signs caused by diseases and the non-specificity of these signs (Mandl et al., 2004; Buckeridge, 2007). In dairy cattle, two data sources were used for syndromic surveillance: mortality and notifications of abortions (Perrin et al., 2010; Brouwer et al., 2012). This reduces the number of detectable signs and, therefore, the number of detectable diseases. Reproduction data have never been used for syndromic surveillance despite their potential interest: they are routinely collected and have been successfully used to build indicators of reproductive disorders associated with infectious diseases (Marce et al., 2009; Santman-Berends et al., 2010; Nusinovici et al., 2012).

The second challenge is to implement statistical methods able to detect small anomalies produced in the early phase of an epidemic while limiting the number of false alarms during the non-epidemic periods (Jefferson et al., 2008). Various statistical methods of detection exist but no single method is ideal for every setting (Unkel et al., 2012). The performance of statistical approaches required a specific evaluation of data under study.

The objectives of this study were to compare a set of reproductive indicators and to implement and evaluate statistical methods to detect the emergence of a disease with known effects on reproduction in dairy cattle. The bluetongue outbreak in France in 2007 was used as a case model. In a first part five indicators of reproduction disorders were built and their ability to show elevations from the expected value during an epidemic was evaluated. Then, five statistical methods for time-series were implemented on the indicator with the highest elevation to evaluate the sensitivity, specificity and timeliness of the alarm system.

MATERIALS AND METHODS

Selection of available data

Four French districts with contrasting between-herd prevalences of bluetongue virus (BTV-8) in 2007 were selected for this study. The between-herd prevalence was calculated as a yearly cumulated prevalence in dairy herds in all French districts. Notifications of dairy herds with at least one infected animal with clinical signs of BTV-8 were used. The four districts selected were: Seine-Maritime (between-herd prevalence = 2%), Haute-Marne (15%), Meuse (31%) and Ardennes (57%).

Reproduction data and milk production data collected in the four selected districts were extracted from the national database managed by INRA (Institut National de la Recherche Agronomique). Only Holstein cows (parous at artificial insemination) present in dairy herds enrolled in the official milk recording scheme were kept. Two databases were extracted: (i) a reproduction database was composed by the following variables: date of artificial insemination (AI), date of calving, animal identification numbers (ID) and herd IDs; and (ii) a milk production database was composed by the following variables: date of test-day, milk production at test-day, date of previous calving, parity, animal and herd IDs. In addition, date of birth and date of death/culling were extracted for each animal. Only data recorded between the 1st of January 2002 and the 31st of December 2010 were selected.

Building time series of indicators of reproduction disorders

Five indicators were evaluated as markers of various reproductive disorders. Three were based on the time interval between the first AI (after calving) and the second AI (i.e., return-to-service) and the time interval between the last AI and calving. The five indicators were defined as follows:

(i) **3-week return-to-service**: interval between first and second AI from 18 to 26 days, could be associated with fertility disorder or early embryonic death.

(ii) **delayed return-to-service**: interval between first and second AI from 27 to 77 days, could be associated with fertility disorder or late embryonic death. The bound 78 days corresponds to 3 times the maximum duration of the estrous cycle (3x26).

(iii) **very late return-to-service:** interval between first and second AI from 78 to 200 days, could be associated with abortion or fetal death.

(iv) **premature calving:** interval between last AI and calving from 230 to 270 days, could be associated with abortion or still birth.

(v) **short gestation**: interval between last AI and calving from 270 to 278 days, could be associated with occurrence of a health disorder in the pregnant animal at the end of gestation.

For each indicator and district, time series of daily rate of occurrence were calculated on a weekly basis. The weekly number of indicators events was divided by the weekly number of animal-days at-risk. For a given day, the number of animals at-risk corresponded to the animals present that day (not dead or culled) with no previous return-to service or calving since last AI and inseminated in the time interval corresponding to the indicator considered.

Comparing the increase in indicator occurrences during the BTV-8 epidemic

The increase of occurrence for the five indicators during the BTV-8 epidemic period was evaluated during the BTV-8 epidemic period in each district. The BTV-8 epidemic period started from the first notification by a veterinarian recorded in each district and ended four months later.

A simple statistical model was implemented to detect abnormal elevations during these periods. The statistical model estimated the expected baseline and the expected upper bound of the random fluctuations in absence of an epidemic. The baseline was modeled by a trigonometric temporal regression (Eq. (1)) with binomial error and logit link. The upper bound of the 95% confidence interval was calculated by Normal approximation (Eq. (2)) using prediction of Eq. (1).

$$\ln\left(\frac{p}{1-p}\right) = \sum_{n=1}^{3} \alpha_n \times \cos\left(\frac{2\pi tn}{52}\right) + \beta_n \times \sin\left(\frac{2\pi tn}{52}\right) \quad (1)$$

$$U = \hat{p} + 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{N}} \quad (2)$$

where, p was the daily rate of a given indicator in a given week t (number of weeks since the 1st of January 2003). The parameters α_n and β_n, were estimated by maximum likelihood using the R glm function (R Core Team, 2012) and observations from 2003 to 2006. U was the upper bound of the 95% confidence interval, $\hat{p}$ was the predicted daily rate of a given indicator in a given week *t* and *N* was the corresponding number of animal-days at-risk.

Two criteria were used to evaluate indicators during the BTV-8 epidemic period: (i) the number of weeks with abnormal elevation (over U) and (ii) the time interval between the first week with abnormal elevation and the first notification by a veterinarian.

Comparison of statistical methods and evaluation of the alarm system to detect the BTV-8 epidemic early

To compare statistical methods, only time series of daily rates of a single indicator were used. This indicator was chosen as the one with the highest number of abnormal elevations during the BTV outbreak. Expected baseline of time series of this indicator in absence of an epidemic was predicted according to five different models: (1) logistic regression with cow-level individual covariates (e.g., parity, milk production), (2) auto-regressive and moving average, (3) exponentially weighted moving average, (4) trigonometric regression over time, (5) Farrington algorithm. Three models were updated each year using data of the last year as additional data for model parameter estimations: auto-regressive and moving average, exponentially weighted moving average and trigonometric regression over time. By definition, the model used by the Farrington method is updated on a weekly basis. To avoid having models account for abnormal rates which occurred during the BTV-8 epidemic, data during this period were replaced by historic observations. Logistic regression was performed on observations between 2003 and 2006 (period without major epidemics). The accuracy of each model was estimated by comparing the mean absolute percentage error (MAPE).

Persistent elevations in observed time series were detected according to the CUSUM method (cumulative sum of differences between expected and observed data; Eq. (3)).

$$S_{t+1} = \max(0, S_t + p_t - k) \qquad (3)$$

Where S_{t+1} and S_t were the CUSUM values on week t+1 and t respectively, p_t is the observed daily rate for the week t, S_0 was equal at zero and *k* was a parameter set by the user. The k value was chosen as the upper bound of the 50% confidence interval (Eq. (2) by replacing 1.96 by 0.67). When S_t was higher than a threshold *h* set by the user an alarm was triggered. A sensitivity analysis was performed on the *h* parameter whose values ranged from 0.005 to 0.05 by 0.0025.

To evaluate the alarm system, three criteria were used: the ability to detect a significant increase in the district during the 2007 bluetongue outbreak, the precocity of the detection (in comparison to BTV notifications of clinical cases) and number of false alarms (in the time period of 3.5 years between 2007 and 2010 without the bluetongue period).

RESULTS

Abnormal elevations of the reproductive indicators during the BTV-8 epidemic

The time series built for each reproductive indicator revealed important differences between the daily rates of indicators before the epidemic (Figs. 1 and 2). Indicators 3-week return-to-service and short gestation had the highest daily rates. The lowest rates were observed for very late return-to-service and premature calving. Modeled baselines were able to reproduce seasonal variability with a MAPE close to zero for all time series (~0.1%). The confidence interval slightly underestimated the observed random fluctuations of time series including only 94.5% of the observations (instead of 97.5%) over all indicators and districts.

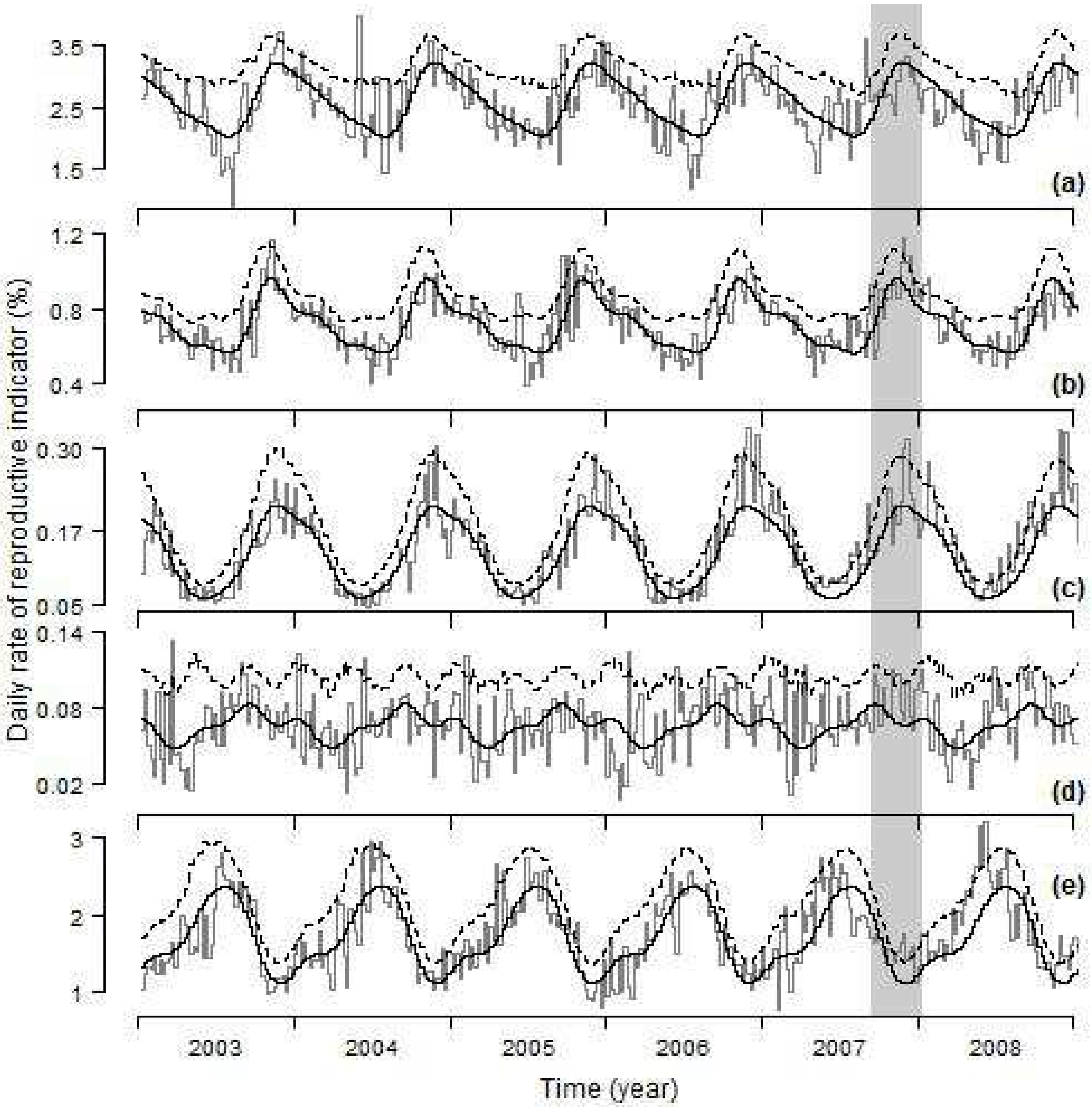


Fig. 1 Observed (grey line), predicted (black line) and upper bound (dashed black line) of daily rates of five reproductive indicators in dairy cows over time in the district Seine-Maritime (France). The five indicators are: 3-weeks return-to-service (a), delayed return-to-service (b), very late return-to-service (c), premature calving (d) and short gestation (e). Daily rates are calculated in Holstein cows of dairy herds enrolled in the official milk recording scheme. The grey area corresponds to the BTV-8 epidemic period in this district where the yearly cumulated prevalence of notified herds was ~2% in 2007

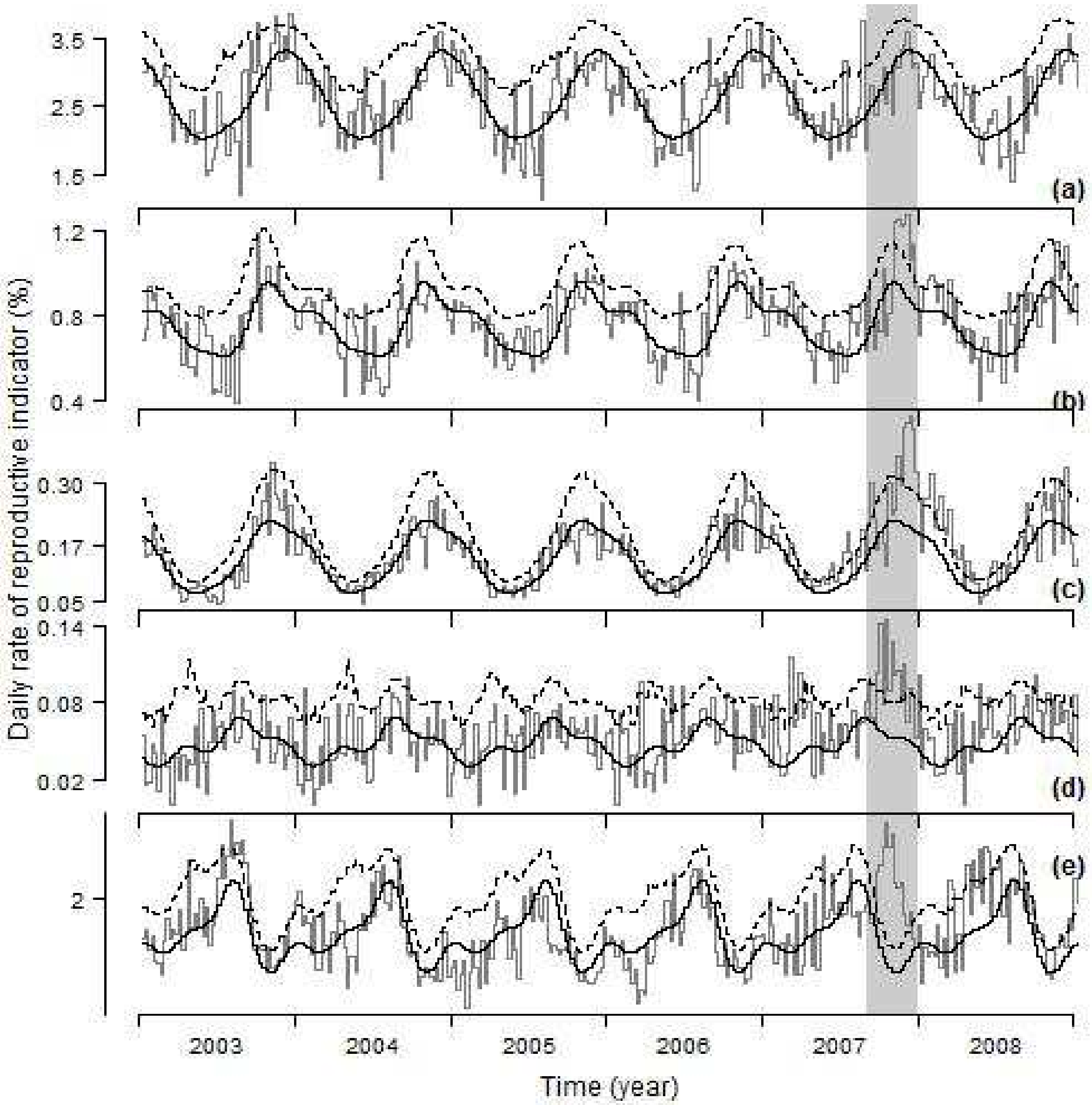


Fig. 2 Observed (grey line), predicted (black line) and upper bound (dashed black line) of daily rates of five reproductive indicators in dairy cows over time in the district Meuse (France). The five indicators are: 3-weeks return-to-service (a), delayed return-to-service (b), very late return-to-service (c), premature calving (d) and short gestation (e). Daily rates are calculated in Holstein cows of dairy herds enrolled in the official milk recording scheme. The grey area corresponds to the BTV-8 epidemic period in this district where the yearly cumulated prevalence of notified herds was ~57% in 2007

The number of weeks with abnormal elevations during the BTV-8 epidemics in each district was high for short gestation and intermediate for premature calving, very late return-to-service and delayed return-to-service (Fig. 3). On the contrary, 3-weeks return-to-service had a number close to zero whatever the district. The number of weeks with abnormal elevations was correlated with the between-herd prevalence in districts: the higher the between-herd prevalence, the higher the number of weeks with abnormal elevations. The median time interval between first notification of a clinical case and first occurrence of an abnormal elevation in a district was: 61, 77, 26 and 19 days for delayed return-to-service, very late return-to-service, premature calving and short gestation, respectively.

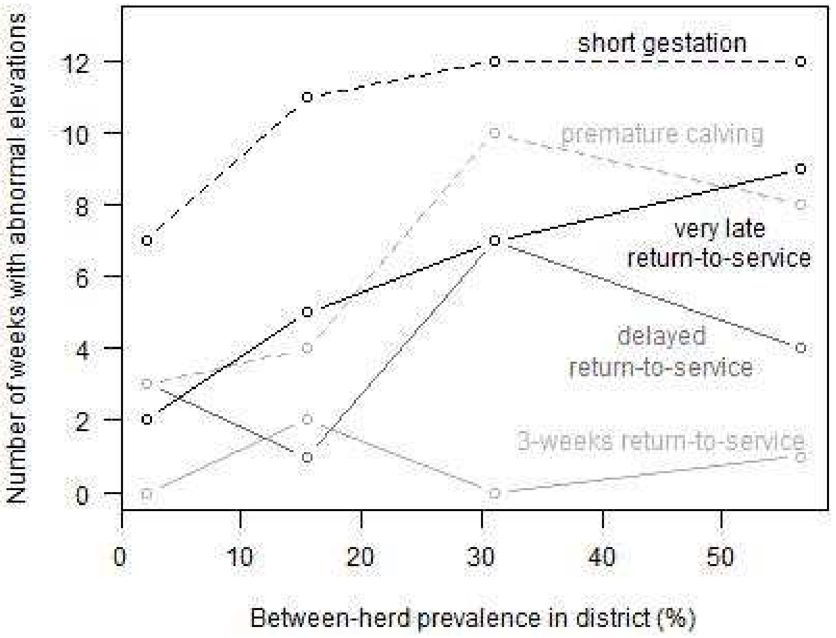


Fig. 3 Number of weeks with abnormal elevations within four months after emergence of the disease in 2007 for five reproductive indicators in dairy cows as a function of between-herd prevalence of notified BTV-8 in four districts in France. The four districts are: Seine-Maritime (between-herd prevalence = 2%), Haute-Marne (15%), Meuse (31%) and Ardennes (57%)

<u>Performance of the statistical methods and alarm system to early detect the BTV-8 epidemic</u>

The indicator short gestation had the highest number of abnormal elevations during the BTV epidemic and was chosen for the comparison of statistical methods to evaluate an alarm system. The five models had similar accuracy with a MAPE which ranged from 13.6% for temporal regression to 16.9% for the Farrington method.

In the three districts with the highest between-herd prevalence, the BTV epidemic was always detected for any *h* value (threshold of the CUSUM to trigger the alarm) lower than 0.04. The higher the h value was, the lower the number of false alarms and the longer the alarm was delayed compared to the first detection of a clinical case of BTV in a district (Fig. 4). Whatever the *h* parameter and statistical model, the BTV epidemic was never detected in the district with the lowest between-herd prevalence.

For the same values of the *h* parameter, three statistical methods had similar numbers of false alarms and delays to trigger an alarm: auto-regressive and moving average, trigonometric regression over time and exponentially weighted moving average. Logistic regression with covariates had a lower number of false alarms but longer delays to trigger an alarm. The Farrington method had long delays to trigger an alarm and a high number of false alarms.

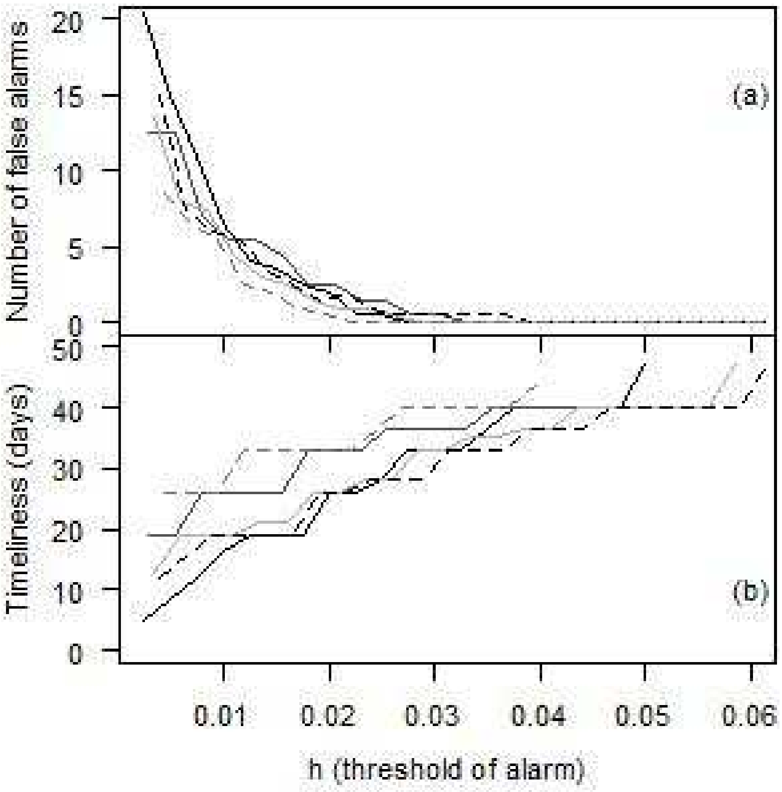


Fig. 4 Number of false alarms (a) and timeliness (b) as a function of the threshold of alarm (*h* parameter) of the CUSUM algoritm applied on times series of an indicator of short gestation in dairy cows modeled by five statistical models. The five models were: auto-regressive and moving average (black line), Farrington algorithm (grey line), trigonometric regression over time (light grey line), exponentially weighted moving average (black dashed line) and logistic regression with cow-level individual covariates (grey dashed line). Curves correspond to the median values observed in four French districts. Timeliness was calculated as the time interval between first notification of BTV-8 in each district and first alarm

DISCUSSION

Four indicators out of five showed detectable elevations in the BTV epidemic period when calculated on a weekly time step at a district level. Elevations lasted longer when the between-herd prevalence in the district increased. Indicators associated with late embryo mortality, foetal loss and abortions (known effects of BTV infection) were elevated during the epidemic, whereas elevations in 3-week return-to-service (due to infertility or early embryo mortality) were not detected. A new indicator, which has not been studied as a possible consequence of BTV, showed the highest number of abnormal elevations: the frequency of short gestations (with pregnancy length within the lowest quartile of the normal distribution). Contrary to the other indicators, the biological mechanism underlying these variations is unknown. If it was due to the effect of an acute infection (with fever) on cows close to parturition, it could be of interest for many infectious diseases. The indicators built here are non-specific and therefore good candidates for syndromic surveillance. Their ability to show (or not) detectable elevations is likely to depend on the disease. It is recommended to use the set of five indicators together to increase the number of detectable diseases with different effects on reproduction.

The alarm system to detect the BTV epidemic from elevations in the frequency of short gestation was effective whatever the statistical model applied to estimate the baseline of time series of daily rates. All systems triggered very few false alarms and showed good timeliness. Three methods resulted in earlier alarms (from 1 to 2 weeks) and corresponded to the three models updated yearly: auto-regressive and moving average, trigonometric regression over time and exponentially weighted moving average. With the indicator short gestation, the choice of a statistical model can be guided by simplicity of modelling rather than by performance of the alarm system.

Performances of the alarm system highly depend on the choice of the threshold value on the CUSUM to trigger an alarm (parameter *h*). As expected, when *h* decreases, the number of false alarm increases and the timeliness of detection is improved (Buckeridge, 2007). A good compromise was obtained for *h*~0.015 with a number of false alarm ranging between 2 and 4 (over 4 years) and timeliness of detection of the BTV epidemic of around 3 weeks. The choice of a threshold depends on the priorities of the decision-makers between specificity of the alarm system and timeliness of an alarm in case of emergence. Sensitivity analysis as performed here can provide them with the necessary information to make a choice. Whatever the statistical models and *h* parameter, the district with the lowest between-herd prevalence was never detected. Here, the signal produced by the BTV epidemic was not high enough to be distinguished with the fluctuation of the times series in the absence of an epidemic, which can be random or due to causes other than a disease.

ACKNOWLEDGEMENTS

This study was funded by ANR (Agence Nationale de la Recherche) in the framework of the European project EMIDA ERA-NET “Early Detection Data”. Data were provided by CTIG (Centre de Traitement de l’Information Génétique) and DGAL (Direction Générale de l’Alimentation).

REFERENCES

Brouwer, H., Bartels, C.J.M., Van Wuijckhuise, L. and Van Schaik, G. (2012). Syndromic surveillance on blood samples of aborting cows for early detection of (emerging) infectious diseases using different statistical methods. In: Society for Veterinary Epidemiology and Preventive Medicine. Proceedings of a meeting held in Glasgow, UK, 28th-30th March 2012, 233-244

Buckeridge, D.L. (2007). Outbreak detection through automated surveillance: A review of the determinants of detection. J. Biomed. Inform. 40, 370-379

Jefferson, H., Dupuy, B., Chaudet, H., Texier, G., Green, A., Barnish, G., Boutin, J.P. and Meynard, J.B. (2008). Evaluation of a syndromic surveillance for the early detection of outbreaks among military personnel in a tropical country. J. Public Health (Oxf.) 30, 375-383

Mandl, K.D., Overhage, J.M., Wagner, M.M., Lober, W.B., Sebastiani, P., Mostashari, F., Pavlin, J.A., Gesteland, P.H., Treadwell, T., Koski, E., Hutwagner, L., Buckeridge, D.L., Aller, R.D. and Grannis, S. (2004). Implementing syndromic surveillance: A practical guide informed by the early experience. J. Am. Med. Inf. Assoc. 11, 141-150

Marce, C., Beaudeau, F., Bareille, N., Seegers, H. and Fourichon, C. (2009). Higher non-return rate associated with Mycobacterium avium subspecies paratuberculosis infection at early stage in Holstein dairy cows. Theriogenology 71, 807-816

Nusinovici, S., Seegers, H., Joly, A., Beaudeau, F. and Fourichon, C. (2012). Increase in the occurrence of abortions associated with exposure to the Bluetongue virus serotype 8 in naïve dairy herds. Theriogenology (in press)

Olson, D.R., Heffernan, R.T., Paladini, M., Konty, K., Weiss, D. and Mostashari, F. (2007). Monitoring the impact of influenza by age: Emergency department fever and respiratory complaint surveillance in New York City. PLos Med. 4, 1349-1361

Perrin, J.B., Ducrot, C., Vinard, J.L., Morignat, E.. Gauffier, A., Calavas, D. and Hendrikx, P. (2010). Using the National Cattle Register to estimate the excess mortality during an epidemic: Application to an outbreak of Bluetongue serotype 8. Epidemics 2, 207-214

R Core Team. (2012). R: A Language and Environment for Statistical Computing.

Santman-Berends, I., Hage, J.J., van Rijn, P.A., Stegeman, J.A. and van Schaik, G. (2010). Bluetongue virus serotype 8 (BTV-8) infection reduces fertility of Dutch dairy cattle and is vertically transmitted to offspring. Theriogenology 74, 1377-1384

Unkel, S., Farrington, C.P., Garthwaite, P.H., Robertson, C. and Andrews, N. (2012). Statistical methods for the prospective detection of infectious disease outbreaks: a review. J. R. Stat. Soc. Ser. A Stat. Soc. 175, 49-82

OPEN SESSION

BAYESIAN EPIDEMIC RISK PREDICTION – KNOWLEDGE TRANSFER AND USABILITY AT ALL LEVELS

C. JEWELL[*], J. BROWN, M. KEELING, L. GREEN AND G. ROBERTS

SUMMARY

Recent advances in Bayesian inference have provided a powerful solution for harnessing epidemic models for decision support during livestock epidemics. However, efficient fitting algorithms for complex epidemic models are difficult to implement. Hand-coded algorithms are required for each application, requiring sophisticated techniques for ensuring reliable estimation of the parameter posterior distributions. Moreover, the complex analysis output may be difficult to interpret, reducing the utility of the approach when fast, clear results are required.

This paper presents an integrated approach to risk prediction, focusing on making Bayesian analysis for epidemics accessible to users at different levels: a web-based GIS interface for the policy maker, an R package for the epidemiologist, and a C++ interface for the statistician. To demonstrate how these levels work in concert, an analysis of the 2001 UK foot and mouth outbreak is presented, extending established models to investigate for the first time the role of pig farms.

INTRODUCTION

The economic, social and welfare implications of epidemic diseases in livestock necessitate swift and decisive control measures implemented on a national scale. In an outbreak situation, disease control authorities are often faced with having to make decisions early on when the dynamics of the epidemic are poorly understood. Although local disease control policy may be implemented effectively by those in the vicinity, coordination of the national scale effort is often hindered by difficulties in sharing and visualising the progress of the epidemic on a national scale.

Mathematical modelling has become a popular method for investigating disease dynamics, characterising the behaviour of the disease process in the population of interest. It has the advantage that disease outbreaks can be simulated in a virtual population, and the effect of varying the model parameters on the outcome of the epidemic can be explored. For example, in UK 2001 foot and mouth disease epidemic, simulations from epidemic models were used

---

[*] Dr Chris Jewell, Infectious Disease Research Centre, Massey University, New Zealand. Email: c.p.jewell@massey.ac.nz

investigate the likely effect of certain control measures on the course of the outbreak (Keeling et al., 2001; Ferguson et al., 2001). A standard setup for an epidemic model is to suppose that an individual (defined as a single epidemiological entity, e.g. an animal, farm, or region) can exist in one of a variety of states, between which the transitions are well defined. For example, the SIR model assumes that individuals begin as susceptible before becoming infected, and are subsequently removed (i.e. recover with solid immunity or die) from the population and play no further part in the epidemic. The parameters of the model typically govern the transition rate from one state to the next, and it is these that may be varied to investigate the behaviour of the model under different conditions (Anderson & May, 1991).

In order to use epidemic models in a predictive sense, however, it is necessary to have reliable values for the parameters. As epidemics are rare events, changes in the pathogen, host, and management factors over time means that suitable parameters for one outbreak do not necessarily apply to the next. For this reason, an approach to estimating such values from epidemiological observations as an outbreak evolves is required.

Recent advances in Bayesian inference have provided a powerful solution for harnessing epidemic models for decision support during livestock epidemics (Jewell et al., 2009a). The approach takes the epidemic case time series as an outcome variable, and known population data (e.g. the locations of farms, the number and species of animal present on each farm) as explanatory variables, to provide estimates of the model parameters together with a rigorous measure of uncertainty. These estimates can then be used to drive the same model forward from the current state of the epidemic to provide quantitative predictions for critical aspects such as the risk to individuals posed by the current state of the epidemic, the likely locations of undetected infections, and identification of high-risk farms. The sequential inclusion of field data as the outbreak unfolds means that risk predictions are updated and refined as the outbreak progresses.

Despite this power of Bayesian inference for epidemics, efficient Markov-Chain-Monte-Carlo (MCMC) algorithms for fitting complex models are difficult to implement. They must be hand-coded for each application, often requiring sophisticated statistical and computational techniques to ensure reliable estimation of the parameter posterior distributions. This is time consuming for the specialist, and may put the methodology out of reach for non-specialist researchers and disease managers. Furthermore, the complex nature of the analysis output may be difficult to interpret, reducing the utility of the approach when concise results are required quickly in an emergency situation.

To address this issue, this paper describes a novel integrated approach to risk prediction, focusing on making Bayesian analysis for epidemics accessible to users at different levels. An R package provides access to a restricted class of epidemic models and is aimed at epidemiologists wishing to analyse epidemic time series on commonly found population data structures. Underpinning the R package is a low level C++ application programming interface, which can also be used for rapid development of new epidemic models, together with MCMC and simulation engines to analyse data. Finally, we show how output from these tools can be visualised using an interactive web-based geographical information system (GIS) oriented towards policy makers during a disease outbreak.

As a demonstration of the framework, the model of Keeling et al. (2001) is extended to include the role of pig farms in the 2001 UK foot and mouth outbreak. This extension presents a surprising challenge for the MCMC fitting algorithm. With the C++ application programming interface (API), a non-centered MCMC algorithm is constructed which is capable of fitting our three-species model, showing how careful tailoring of the MCMC is required for reliable inference. Embedding this code into our R package, the relative importance of the three species in the epidemic is investigated, showing that, from a farm level perspective, pigs were less infective and less susceptible to FMDV Type-O Pan Asia than cattle and sheep. Finally, the web-based GIS is used to summarise our results and identify high-risk farms during the outbreak.

MATERIALS AND METHODS

Approach to risk prediction

In order to characterise the epidemic process in the study population, a stochastic 'SINR' epidemic model is specified in which farms are taken to be individual epidemiological units (Jewell et al., 2009b). Farms are considered to start out as susceptible, then become infected, notified (i.e. the infection detected and notified to the disease authorities), and finally removed (i.e. depopulated) playing no further part in the epidemic. The progression from susceptible to infected is governed by the 'transmission process', and from infected to notified by the 'detection process'. These two processes are both stochastic, are conditionally independent (i.e. conditional on the infection events), and are described below. The progression from notified to removed is considered to be deterministic, since from a statistical point of view both the notification and removal events are observed, whereas the infection event is typically unobserved.

In contrast to much of the literature (see for example Tildesley et al. (2008)), here the term 'model' is defined to refer specifically to the mathematical conceptualisation of the epidemic process. Statistical inference may be performed on the model, inferring the rates of progression between the four states given epidemic data. Alternatively, an epidemic may be simulated from the model given estimates of the rates of progression. Respectively, this is analogous to obtaining estimates of coefficients of a linear model given explanatory and outcome variables, and using the model to predict an outcome given a set of coefficients and explanatory variables.

Transmission process

Farms are assumed to become infected according to a continuous time-inhomogeneous Poisson process where the rate of infection $\lambda_j(t)$ on an individual farm j at time t is the sum of the rate of infection from all infected and notified farms in the population at time t (see Eq. (1))

$$\lambda_j(t) = \epsilon(t) + \left[\sum_{i \in \mathcal{I}(t)} \beta_{ij} h(t - I_i \vee 0) + \sum_{i \in \mathcal{N}(t)} \beta^{\star}_{ij}\, h(t - I_i \vee 0)\right] \quad (1)$$

where $\mathcal{I}(t)$ and $\mathcal{N}(t)$ are the sets of infected and notified individuals at time t, I_i the infection time of individual i. β_{ij} and $\beta^{\star}_{ij}$ are the pairwise rates of infection between an infected or notified (respectively) farm i and a susceptible farm j, with $h(\cdot)$ describing how an infected farm's infectiousness changes with time and may be specified according to the particular application of the model. $\epsilon(t)$ denotes a (possibly time-dependent) background rate of infection that is due to exogenous sources of infection such as wildlife.

The pairwise transmission rates β_{ij} and $\beta_{ij}^{\star}$ are specified in terms of a gravity model, dependent on the infectivity $q(i;\boldsymbol{\xi},\boldsymbol{\psi})$ of farm i, the susceptibility $(j;\boldsymbol{\zeta},\boldsymbol{\phi})$ of farm j, as well as a function of distance between them $K(i,j;\boldsymbol{\delta})$, see Eq. (2)

$$\begin{gathered}\beta_{ij} = \gamma_1 q(i;\boldsymbol{\xi},\boldsymbol{\psi}) s(j;\boldsymbol{\zeta},\boldsymbol{\phi}) K(i,j;\boldsymbol{\delta}), i \in \mathcal{I}, j \in \mathcal{S} \\ \beta_{ij}^{\star} = \gamma_2 \gamma_1 \beta_{ij}, i \in \mathcal{N}, j \in \mathcal{S} \end{gathered} \tag{2}$$

γ_1 is interpreted as the baseline rate of infection between the farms, whilst γ_2 measures the effect of placing statutory restrictions on the infected farm as a result of the disease notification process. The exact form of the infectivity, susceptibility, and distance functions are application dependent.

Disease detection process

To model the period of time for which an individual farm spends in the infected state, an application-dependent probability function is used. Typically, a Gamma distribution is used such that

$$N_i - I_i = d_i \sim \text{Gamma}(a,b)$$

In principle, however, a wide range of probability distributions can be used providing $0 \leq d_i \leq \infty$.

Since the notification and removal events are both observed, they are treated as deterministic and no probability model is applied.

Statistical inference and risk prediction

Having described a model for an epidemic process, it is necessary to consider which components are known (or observed) in advance of, or during, a disease outbreak and which are inherently unobserved. The static population data, that is the locations of farms together with information such as their size and management types, are known in advance of an epidemic and constitute covariate information. The timings, and identities, of notification and removal events are directly recorded throughout the epidemic, and may be regarded in statistical terminology as the observations.

The remaining components of the model, the parameters $\boldsymbol{\theta} = \{\gamma_1, \gamma_2, \boldsymbol{\xi}, \boldsymbol{\psi}, \boldsymbol{\zeta}, \boldsymbol{\psi}, \boldsymbol{\delta}, a, b\}$ and times of the infection events, are typically unknown and must be estimated from the observed data. The inherent inability to observe infection events gives rise to the possibility of undetected infections present in the population at any point during the epidemic. In order to obtain unbiased estimates of the model parameters, it is necessary to account for both the censored data and undetected infections.

Performing inference in a Bayesian paradigm offers several advantages over a classical (frequentist) approach. During the early stages of an epidemic when there are few case detections on which to make inference, it is necessary to fall back on past experience of disease outbreaks, in the form of *prior* information. Prior information on the model parameters is encoded in the form of probability distributions where the variance of a distribution is directly related to the prior uncertainty about the value of the parameter. Using a probability model, the joint probability distribution over all the parameters is then updated by the observations, giving the *posterior* probability distribution (Bernardo & Smith, 2000). As with the prior, the posterior distribution mathematically encodes the *a posteriori* statistical uncertainty in a form which is easily fed forward and reflected in the final risk prediction results.

In addition, using a Bayesian approach to risk prediction provides a natural way to account for the censored data, treating unobserved infection times as parameters in the model. The power of transdimensional MCMC techniques provides a way to fit the Bayesian model to data, even for large models with many parameters and missing infection times. Moreover, it is able to attribute probabilities to each presumed-susceptible farm that it is, in fact, infected but currently undetected.

Once the joint posterior distribution over the parameters and missing data has been estimated, it may be fed forward into a stochastic simulation. The simulation is started from the current state of the epidemic (bearing in mind the posterior distribution encompasses the presence of undetected infections), and runs forward in time to provide the Bayesian *predictive probability distribution* of the risk posed to the population by the current state of the epidemic. Importantly, the predictive distribution includes a rigorous measure of statistical uncertainty, meaning the associated risk predictions will therefore not be under or over confident.

Since the purpose of this paper is to present these methods in an approachable form for the general research user, details of the fitting and simulation algorithms are not presented here, but are available in Jewell et al. (2009b).

R package “BERP”

The R package “BERP: Bayesian Epidemic Risk Prediction” is aimed at providing Bayesian epidemic risk prediction for common situations. It is provided under the GPLv3 open source software license, and available publicly on the internet via a source code repository (https://github.com/chrism0dwk/infer). BERP encapsulates individual-level and spatially aggregated epidemic models, and is intended as a research tool for epidemiologists. It removes the need for specialist knowledge of MCMC, and presents a familiar R frontend interface for managing data and running the algorithms. Full access is given to the algorithm outputs, allowing researchers to fully examine and critique their results.

The basic workflow for BERP is as follows: A model for the epidemic is first selected based on the structure of the population covariate dataset (farm locations, number of animals, etc.), and the epidemic observations (times of case detections, farm depopulations, etc.). Currently, BERP provides an individual level model and an area-aggregated model; this paper uses the former as the example of choice. The model is supplied with covariate information as a standard R `data.frame` object, including farm identification (e.g. county parish holding number), the numbers of animals in up to three production types present on the farm, and the farm coordinates. Coordinates are specified using a transverse Mercator projection, no automatic conversion of coordinate systems is currently implemented. The epidemic data are supplied to the model as a data frame with columns for farm identification, best guess infection time (noting that these will be inferred by the fitting process), notification time, and removal time. A further column specifies the infection type, enabling the model to differentiate between different types of removed premises (e.g. known infection, dangerous contact, contiguous premises, etc.). The distributional forms for the priors are currently fixed for each parameter in the model, but their hyperparameters may be supplied as an R `list` object to the constructor. Finally, the observation time (i.e. the time during the epidemic at which the analysis is being done) is specified. For example

```
  > library(BERP)
  > theModel <- SpatPointSINR(thePopulation,theEpidemic,prior,
obsTime)
```

A generic interface is then used to fit the model, and obtain an estimate of the joint posterior distribution. The package automatically selects the most suitable fitting algorithm for the model based on the model type, thus shielding users from the unnecessary complication of matching algorithms to models manually. This allows optimisation of the algorithm to specific models, providing maximal computation speed and efficiency.

For the SpatPointSINR model, the `berp.fit()` command selects a transdimensional MCMC fitting algorithm for the model and returns a `Posterior` object. For example

```
  > thePosterior <- berp.fit(theModel)
```

Note that this example omits certain arguments that may be supplied to `berp.fit()` to control the behaviour of the (MCMC) fitting algorithm used. These are documented in the function online help.

Since the size of the posterior dataset obtained from the fitting algorithm may exceed the amount of available computer memory, the `Posterior` class represents an opaque data structure stored in a temporary file on disc using the HDF5 format (The HDF Group, 2012). The numerical estimates of the posterior are obtained using the usual subscript (‘[]’) and field (‘$’) operators in R, with the data being transparently loaded when required. This gives the user great flexibility to choose the number of burn-in MCMC iterations, as well as thin the posterior draws easily. For example

```
  > delta <-
thePosterior$parameters$delta[seq(30000,100000,by=100)]
```

returns a vector containing the marginal posterior realisations for the δ parameter minus 30000 burn in iterations, and thinned by taking every 100th iteration.

BERP also provides a suite of simulation engines to generate virtual epidemics given the model class. Again, the backend simulation algorithm is chosen based on the model class. For example

```
  > theSim <- berp.sim(theModel, myParameters)
```

returns a `data.frame` containing the identifiers, and infection, notification, and removal times of each infected individual in the simulated SINR epidemic described above, with the parameters contained in the `list` object `myParameters`. Importantly, the simulation algorithm begins from the current state of the epidemic – that is it simulates forward from the observation time applied to the model.

The `Posterior` object may be browsed using R’s `names()` function, and the auxiliary method `summary()` returns a table summarising each marginal posterior distribution. In addition, the `write.posterior()` and `read.posterior()` functions provide the user with the means to save the posterior on permanent storage and to re-access it for further analysis.

C++ API "InFER"

Underlying BERP is a low-level C++ API designed to aid rapid development and customisation of reversible jump MCMC algorithms for fitting Bayesian epidemic models. It is distributed under the GPLv3 open source software licence (https://github.com/chrism0dwk/infer). The API provides C++ classes for the epidemic model with associated likelihood and priors, and a hierarchical MCMC object factory for creating MCMC update steps for the parameters for which inference is required.

The `mcmc` object factory provides several generic MCMC update algorithms, such as single site random walk, and adaptive multisite logarithmic random walk flavours of Metropolis Hastings. An extensible framework is supplied, allowing the user to create additional custom updaters for application-specific tuning. Importantly, the MCMC update algorithms assume no knowledge of how the likelihood is calculated, delegating instead to the `Likelihood` class. Using C++'s class inheritance, this allows specialisation of `Likelihood` to the specific application and available hardware. For example, the `SpatPointSINR` R class makes use of a specialised C++ class that incorporates the SINR model, and utilised the power of a general purpose graphic processing unit (GPGPU) to accelerate the calculation of the likelihood function. Using a NVIDIA Tesla C2075 card provides a 250-fold speedup compared to a single modern central processing unit (CPU).

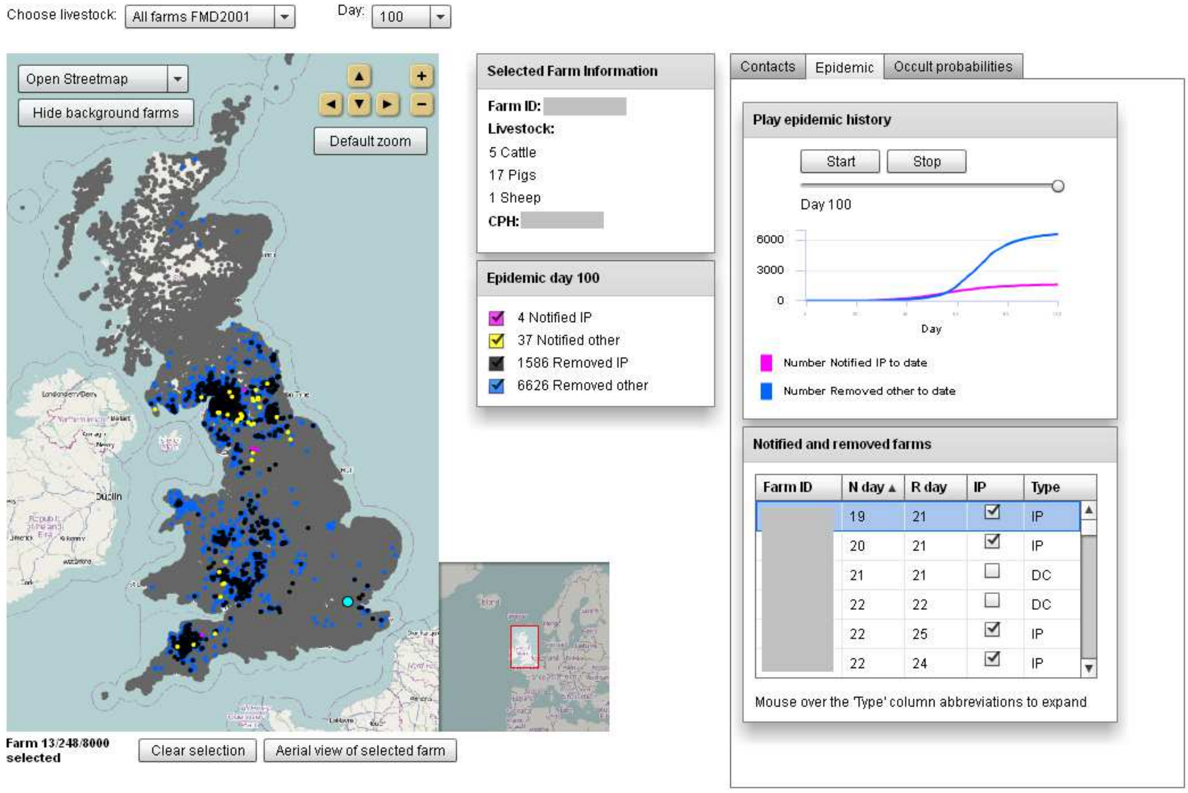
Though only a brief overview is provided here, the purpose of the C++ API is to add flexibility to BERP, such that advanced users may tailor the backend algorithms to new model situations, and novel Bayesian fitting algorithms. It is hoped that this will provide a platform for community development of BERP, increasing its utility in the future.

GIS Web application

The interactive GIS web application provides a graphical user interface (GUI) and visualisation modules to allow policy makers and other end-users to view the epidemic progress and summarised results of the Bayesian analysis in an intuitively understandable format. This tool is being described in detail elsewhere; a brief overview is given here.

The web application has been developed in Adobe ActionScript and runs in any browser with the freely available Adobe Flash Player version 10 or above plug in. The software can be used to visualise the three categories of data described above – covariate data, which describes an individual's attributes and remains constant throughout the epidemic, epidemiological data, which describes the disease states of individual farms and changes as the epidemic progresses, and unobserved data, which is imputed by statistical inference using the other tools in the framework described here. These three data categories are visualised by selecting tabs on the GUI and the selected data is displayed on a zoomable, pannable map as well as shown in tabular and/or graphical form. A screenshot of the web application displaying the FMD2001 epidemic on day 100 is shown in Fig. 1. Interactive elements of the application allow an animation of the epidemic so far to be played and paused (top right) and further information about individual premises can be viewed by clicking on the map or on a table row to select. In Fig.1, the first notified farm has been selected.

The GIS web application allows field workers and policy makers to identify high-risk farms during the outbreak from a remote location and to download and visualise inferred data as well as previously recorded information about those farms and their contact networks.



Choose livestock: All farms FMD2001

Day: 100

Selected Farm Information

Farm ID: [redacted]

Livestock:

5 Cattle

17 Pigs

1 Sheep

CPH: [redacted]

Epidemic day 100

☑ 4 Notified IP

☑ 37 Notified other

☑ 1586 Removed IP

☑ 6626 Removed other

Farm 13/248/8000 selected

Play epidemic history

Day 100

Notified and removed farms

Farm ID	N day ▲	R day	IP	Type
[redacted]	19	21	☑	IP
[redacted]	20	21	☑	IP
[redacted]	21	21	☐	DC
[redacted]	22	22	☐	DC
[redacted]	22	25	☑	IP
[redacted]	22	24	☑	IP

Mouse over the 'Type' column abbreviations to expand

Fig. 1 Screenshot of the GIS web application displaying the FMD2001 epidemic on day 100. The first notified farm has been selected by clicking on the top table row. Farm identities have been obscured to retain confidentiality

CASE STUDY

To demonstrate the framework, the 2001 UK foot and mouth disease outbreak dataset (FMD2001) is used (National Audit Office, 2002). This dataset has been well studied, and is described in detail elsewhere (Kao, 2002). Briefly, the outbreak began with the first case detection on the 19th February, and is regarded to have ended with the last cull some 316 days later on the 1st January 2002. During the epidemic, some 2026 farms were confirmed as infected premises (IP), with a further 7457 farms culled, classed mostly as Contiguous Premises (CP) or Dangerous Contacts (DC). A national movement ban was implemented on the 23rd of February, by which time the disease had been disseminated widely via a chain of livestock markets.

Despite the scrutiny this dataset has received, FMD2001 still presents a convenient example of a large outbreak in a well-characterised population, with many outstanding questions remaining to be asked of it. Previous models of the epidemic have focused solely on the role of cattle and sheep, since these were by far the most numerous species affected, despite cases of the disease in other species occurring (Keeling et al., 2001; Ster & Ferguson, 2007) . In particular, pigs may be important hosts for certain strains of the disease, although in the case of FMD2001 they were present on only 3% of the farm population and just 10 of the IPs. Nevertheless, pigs are known to be susceptible to FMD, and in Europe pig farming represents a significant sector of the livestock industry. It is therefore seen as important to include them in predictive models of the disease process, remembering that in the early stages of an outbreak there may be considerable uncertainty as to the pathogen strain's species predilection. Defining the 19th of February as Day 19, the outbreak is analysed at 4 time points, being Days 33, 50, 100, and 325, as shown in Table 1, classifying CP and DC farms as prematurely culled (PC). Although only a few time points are presented here as examples, it is expected that the analysis would be performed daily during an outbreak.

Table 1. Numbers of infected premises (IPs) and prematurely culled premises (PCs) at each analysis time

Time (days)	Number of IPs	Number of PCs
33	96	110
50	547	632
100	1590	6663
325	2026	8237

The Model

The cattle/sheep model of Keeling et al. (2001) is extended to characterise the disease transmission rate between pairs of infected and susceptible farms in terms of the number of cattle, sheep, and pigs present on each farm. The pairwise transmission rates are defined in Eq. (3).

$$\beta_{ij} = \gamma_1 \left(c_i^{\psi_c} + \xi_p p_i^{\psi_p} + \xi_s s_i^{\psi_s}\right)\left(c_j^{\phi_c} + \zeta_p p_j^{\phi_p} + \zeta_s s_j^{\phi_s}\right)\frac{\delta}{\delta^2 + \rho_{ij}}, i \in \mathcal{I}, j \in \mathcal{S}$$

$$\beta^{\star}_{ij} = \gamma_2 \beta_{ij}, i \in \mathcal{N}, j \in \mathcal{S} \tag{3}$$

where c_k, p_k, and s_k are the number of cattle, pigs, and sheep present on farm *k*. ρ_{ij} represents the distance between farms i and j in kilometres within a one-sided Cauchy distance kernel with decay parameter δ. The time dependent infectivity function is specified such that a farm undergoes a 4 day incubation period before becoming infectious. In other words

$$h(t) = \begin{cases} 1, t \geq 4 \text{ days} \\ 0, \text{otherwise} \end{cases}$$

The movement ban imposed on the 23rd of February (Day 23) is allowed for by considering different background rates of infection before and after the implementation of movement restrictions. Thus,

$$\epsilon(t) = \begin{cases} \epsilon_1, t < 23\text{rd February} \\ \epsilon_1 \epsilon_2, \text{otherwise} \end{cases}$$

with ϵ_1 representing a time-independent background transmission rate, and ϵ_2 measuring the effect of movement restrictions. This effectively treats the initial 'silent spread' phase of the epidemic up to the 23rd of February as a spatial Poisson point process with uniform intensity over the whole country. The infection to notification time is modelled using a Gamma(a,b) distribution, with a=4 and b=0.5 fixed to give a median of 7.3 days, with 5% and 95% quantiles at 2.7 and 15.5 days respectively.

Priors are assigned to each transmission parameter according to Table 2. Gamma distributions represent appropriate choices for rate parameters since they have support on the interval $(0, \infty)$. In this analysis, the prior means and variances are chosen based on poor prior knowledge of the disease process. For the power parameters $\boldsymbol{\psi}$ and $\boldsymbol{\phi}$, Beta(α, r) distributions, having support on (0,1), are chosen since it is expected that infectivity and susceptibility should increase sublinearly with the number of animals present on a farm. For all power parameters, $\alpha = r = 15$ reflecting a moderate belief that infectivity and susceptibility of individual species should be proportional to the square root of the number of animals.

Table 2. Prior distributions for FMD2001 transmission parameters

Parameter	Prior distribution
ϵ_1	Gamma(5e-5,1)
ϵ_2	Gamma(1,1)
γ_1	Gamma(1,1)
γ_2	Gamma(2,4)
$\xi_p, \xi_s, \zeta_p, \zeta_s$	Gamma(1,1)
$\psi_c, \psi_p, \psi_s, \phi_c, \phi_p, \phi_s$	Gamma(15,15)
δ	Gamma(1,1)

Inference

Using BERP to perform the statistical analysis, the model is created using the `SpatPointSINR()` constructor function, which is passed to the MCMC backend fitting algorithm. The addition of pigs to the original Keeling et al. (2001) model presents a surprising challenge for implementing MCMC efficiently. However, the extensible nature of the C++ API allows a custom Metropolis Hastings update step to be written, incorporating a non-centred parameterisation technique for the $\boldsymbol{\xi}$ and $\boldsymbol{\zeta}$ parameters. This results in a vastly improved algorithm (results not shown), which is implemented in BERP.

For each time point, two MCMC chains were run for 100,000 iterations, with 30,000 iterations taken as burn in. The resulting iterations were used to calculate kernel density estimates of the marginal posterior distributions of the parameters and mean infection to notification time for all known infected individuals at the respective time points. A defining feature of the transdimensional MCMC is that it directly assigns to each presumed susceptible individual a probability that it is in fact an undetected infection. These probabilities may then be plotted on a map to guide surveillance teams to farms likely to have been infected.

RESULTS

The MCMC algorithm provides an estimate of the joint posterior distribution for the FMD2001 model, encompassing the transmission parameters, infection times, and undetected infections. Kernel density estimates of the marginal posterior distributions for the parameters are shown in Figs. 2 to 4. For all parameters, the posterior precision increases as more data is obtained from the field. This is reflected in the way the posterior departs from the prior with increasing time. The marginal posteriors for ϵ_2 and γ_2 (Fig. 2) respectively indicate the effect of the movement ban and farm-level restrictions in response to disease detection.

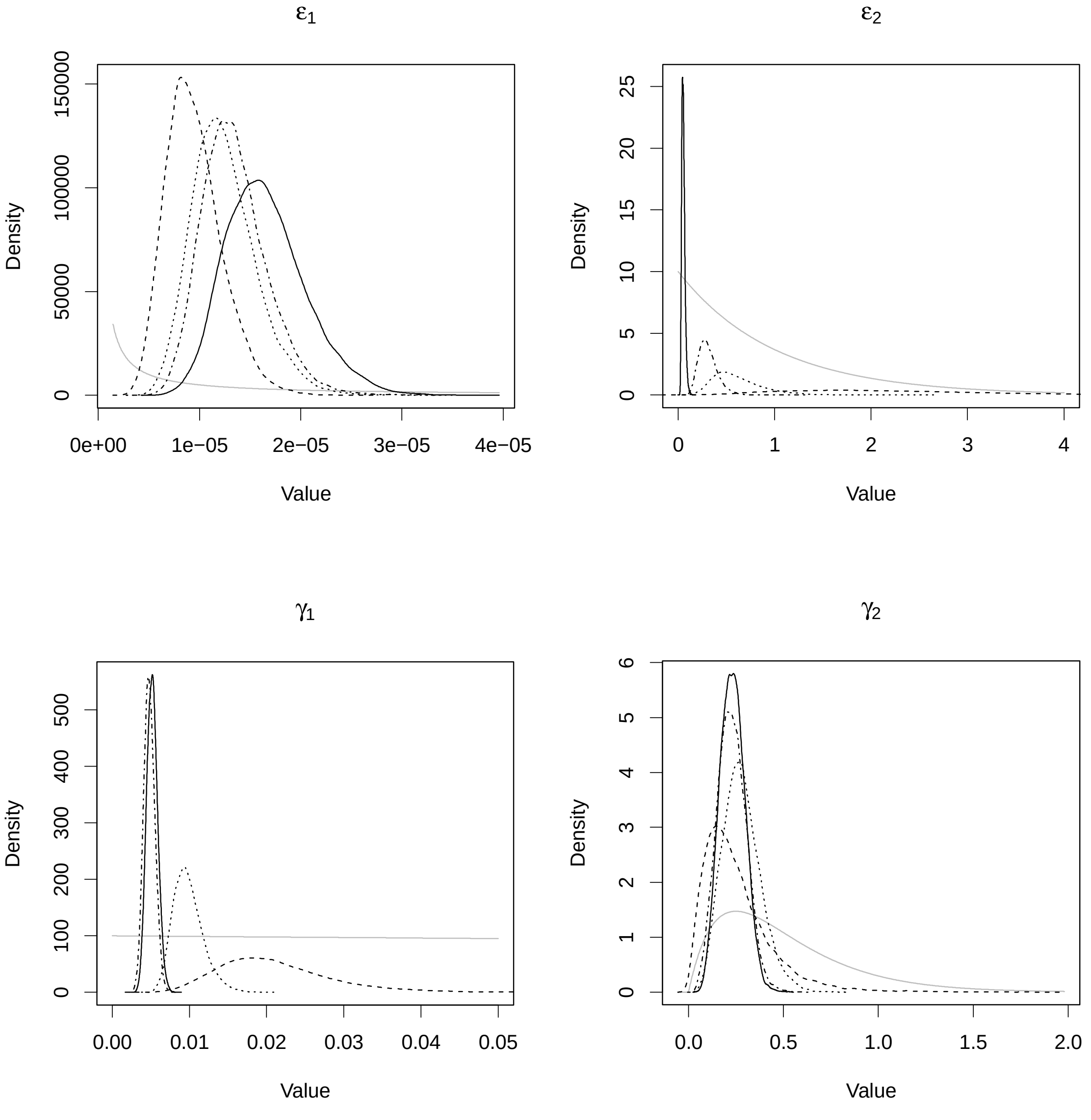


Fig.2 Posterior density estimates for the background transmission parameters ϵ_1 and ϵ_2, as well as γ_1 and γ_2. Grey line ≡ prior; Dashed line ≡ Day 33; Dotted line ≡ Day 50; Dot-Dashed line ≡ Day 100; Solid line ≡ Day 325

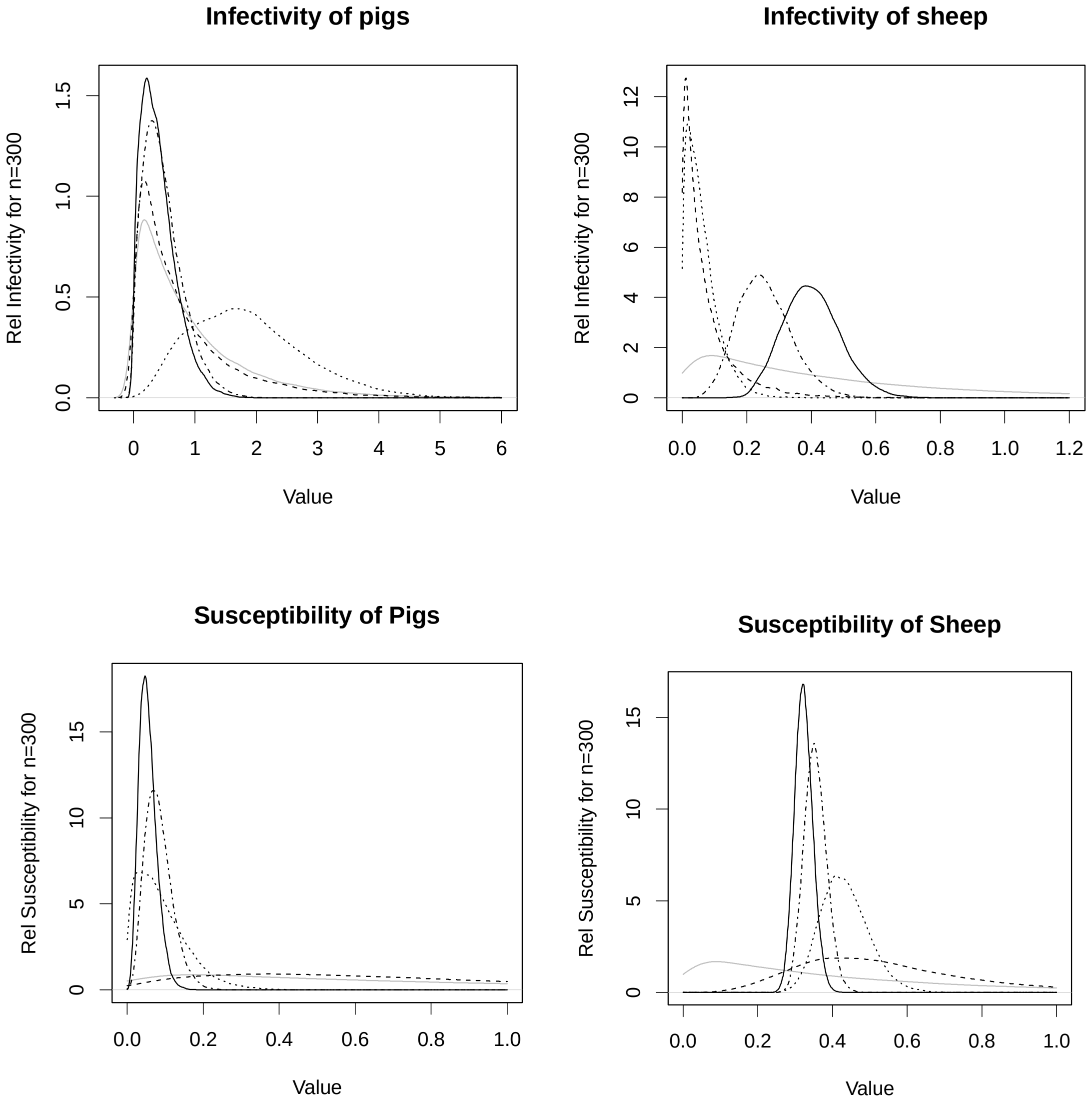


Fig. 3 Marginal posterior probability density estimates for the infectivity and susceptibility of pig herds and sheep flocks relative to cattle herds. Herd/flock size is 300.
Grey line ≡ prior; Dashed line ≡ Day 33; Dotted line ≡ Day 50;
Dot-Dashed line ≡ Day 100; Solid line ≡ Day 325

The non-linear model for infectivity and susceptibility presents a challenge for interpreting the relative importance of each species in the epidemic. Fig. 3 presents the estimates of the marginal posterior distributions for herd/flock-level pig and sheep infectivity and susceptibility relative to cattle, and assumes single-species herds of 300 animals. For pigs, the posterior mass is concentrated on values less than 1 for both infectivity and susceptibility (medians 0.35 and 0.05 respectively for Day 325), with the posterior precision increasing as the epidemic progresses. Sheep appear consistently less infectious than cattle (median 0.40 for Day 325), though interestingly the posterior precision decreases as more data is obtained. Similarly, the model estimates sheep to be less susceptible than cattle (median 0.32 for Day 325). To assess how infectivity and susceptibility vary with herd/flock size, Fig. 4 shows the median marginal posterior infectivity ($q(i;\boldsymbol{\xi},\boldsymbol{\psi})$) and susceptibility ($s(j;\boldsymbol{\zeta},\boldsymbol{\phi})$) as a function of number of animals. Cattle are consistently the most infectious and susceptible species for all her sizes. Pigs are approximately as infectious as sheep for all practical herd/flock sizes, though are estimated to be extremely unsusceptible to FMD.

The occult probabilities at each time point are shown in Fig. 5. The high occult probabilities at Days 33 and 50 reflect both the fast moving nature of the epidemic and the abundance of still susceptible individuals. In comparison, the occult probabilities are lower for Day 100 reflecting the well-known slow progression of the epidemic "tail" (Jewell et al. (2009a). Occult predictions for Day 325 do not appear on the map, since each farm has a probability less than 0.5%.

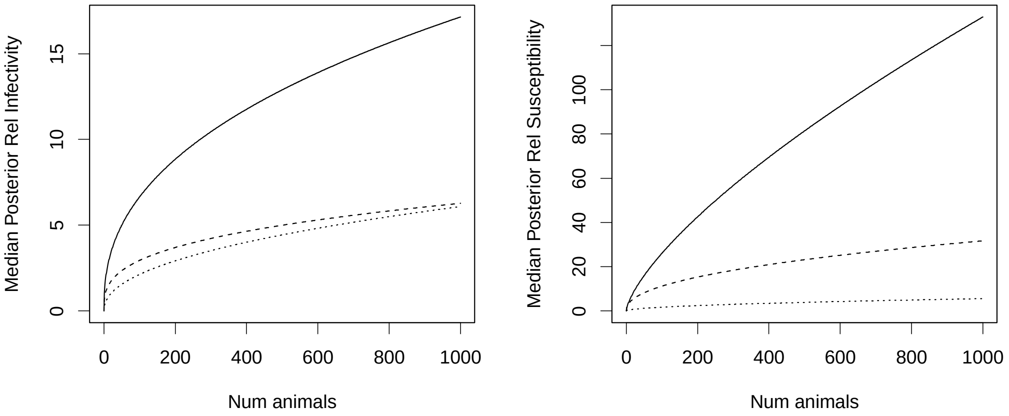


Fig. 4 Median posterior infectivity and susceptibility as a function of number of animals. Solid line ≡ cattle, dashed line ≡ sheep, dotted line ≡ pigs

.

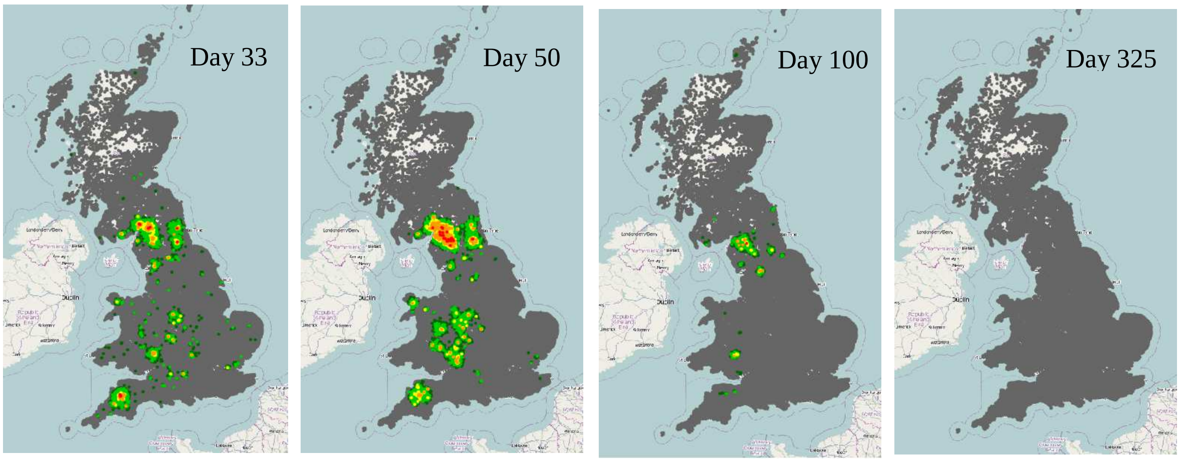


Fig. 5 Visualisation of occult infections during the FMD2001 epidemic using the GIS web application. Premises with probability of being an occult >0.35 (red), 0.1-0.35 (orange), 0.05-0.1 (yellow), 0.01-0.05 (light green), 0.005-0.01 (dark green) are shown for days 33 (top left), 50 (top right), 100 (bottom left) and 325 (bottom right) of the epidemic

DISCUSSION

This paper presents an approach to real time statistical analysis and risk prediction for epidemics that provides entry points for researchers and policy makers with differing levels of statistical expertise. To demonstrate its utility, the example of FMD2001 shows how the C++ API, the R package BERP, and the GIS web interface work together to provide an insight into the dynamics of an unfolding epidemic. Centrally, the Bayesian paradigm provides a formal framework for learning about a disease process, with prior knowledge informing predictions at the very beginning of an outbreak, but being updated swiftly as data arrive from the field. Learning quickly about the dynamics of a new outbreak is a critical feature of effective control, particularly for a disease such as FMD where species predilection may change with pathogen strain evolution. For this reason, being able to communicate the results of a complex high-dimensional statistical analysis to non-experts in an outbreak setting is crucial to its utility.

For the first time, this paper presents an analysis of FMD2001 incorporating the presence of pigs in the national livestock population. The posterior density estimates for ϵ_2 and γ_2 may be regarded as direct measures of control policy efficacy. The former shows a marked decrease in the background infection rate after the movement ban was implemented, preventing the spatially disjointed ‘spark’ infections that characterised the first 14 days of the outbreak. The latter parameter confirms that the restrictions imposed on detected cases were effective at reducing the infectiousness of an IP.

The non-linearity of the infectivity and susceptibility functions for farms complicates the interpretation of the $\boldsymbol{\xi}$ and $\boldsymbol{\zeta}$ parameters. For linear models, for example Ster & Ferguson (2007), in which the powers are implicitly set to unity, these parameters may be interpreted as the relative infectivity and susceptibility of a single sheep or pig relative to a cow. However, in the non-linear case, it may be that for small numbers of animals, the relative infectivities and susceptibilities might be reversed compared to larger herds or flocks. For this reason, it is preferred either to fix our herd/flock size and examine the statistical uncertainty present (Fig. 3), or to take a summary measure of the posterior distribution and investigate the effect of herd size (Fig. 4). Of course, the interactive nature of computer visualisation can greatly help in exploring such multidimensional functions, aiding interpretability.

These results for the FMD2001 example reflect the low numbers of pig farms infected during the outbreak, an unsurprising conclusion given a decade of hindsight. However, as host populations, pathogen strains, and the environment in which they co-exist change over time, a new disease outbreak may not necessarily behave according to precedence. This statistical analysis can, therefore, provide an early insight into the national level dynamics of an unfolding epidemic, enabling control policy to be carefully tailored to the particular disease outbreak setting. Moreover, this information is automatically fed into risk prediction, such as the occult probability maps shown in Fig. 5. In a real time setting, we conjecture that these maps may be useful for targeting disease surveillance to likely undetected infections. An important feature is that they show the probability of a farm being infected, rather than a definite prediction of whether it is infected or not. Farms may then be ranked in order of occult probability, and therefore targeted in an ordered fashion (Jewell & Roberts, 2012).

This paper presents a selection of results from the FMD2001 analysis, though a wide range of further results are available. The numerical estimate of the posterior allows the calculation of predictive quantities of interest, such as the individual-level reproduction number, R_i. In particular, BERP provides the posterior predictive distribution of the epidemic by simulating forward from the current time, using the posterior distribution as parameter estimates together with the predicted locations of infectives. A current limitation in the accuracy of the predictive distribution is due to the provided model classes not explicitly incorporating the process of identifying dangerous contacts, and other active processes that control authorities implement in addition to detection and culling of cases. Therefore, whereas such a model could be readily applied to a situation in which disease control is based purely on reacting to the occurrence of clinical disease, for outbreak such as FMD2001, the predictive simulation does not incorporate the effect of on-going premature culling (e.g. of dangerous contacts, contiguous premises, etc.). The effect of this is such that the predicted future extent of the epidemic necessarily represents a worst case scenario. However, an active area of research is to provide BERP model classes that explicitly measure these processes, and can therefore reflect their effect in predictive simulations.

Software development

Although early in its development, this software suite provides a readily extensible framework for a wide variety of diseases in many populations. BERP currently provides models for individual level SINR epidemic models (such as for the FMD2001 example) and areal metapopulation SIR model (see for example Minh et al. (2011)). As with any statistical technique, the choice of model is strongly influenced by the aspects of the data the analyst is interested in. In terms of focusing purely on prediction, it may be the final risk maps that are of sole interest. However, as the FMD2001 example shows, a finer level of detail is available by considering the marginal posterior estimates themselves shedding light on the global effects of aspects such as species and disease control measures. For this reason, we envisage the on-going development of more model classes to encompass the requirements of epidemic risk prediction in other settings.

A primary design decision in BERP's development was to focus on simplicity of use, and computational efficiency and speed for specific model classes, rather than to provide a high degree of flexibility to the user. The source code is provided as an open source project under the GPLv3 software license, which is intended to stimulate collaboration in developing BERP further, providing a focus of motivation to develop statistical methodology and software for epidemic risk prediction in tandem.

ACKNOWLEDGEMENTS

We gratefully acknowledge Kate Sharpe and Francesca Gauntlett at AHVLA for their invaluable discussions on what disease outbreak administration requires from quantitative epidemic risk prediction.

REFERENCES

Anderson, R.M. and May, R.M. (1991). Infectious diseases of humans: dynamics and control. Oxford University Press, Oxford

Bernardo, J.M. and Smith, A.F.M. (2000). Bayesian Theory. Wiley

Ferguson, N.M., Donnelly, C.A. and Anderson, R.M. (2001). Transmission intensity and impact of control policies on the foot and mouth epidemic in Great Britain. Nature 413, 542-548

The HDF Group. Hierarchical data format version 5, 2000-2010. http://www.hdfgroup.org/HDF5 (Accessed on the 12th of January, 2013)

Jewell, C.P., Kypraios, T., Neal, P.D. and Roberts, G.O. (2009a). Bayesian methodology for epidemics in complex stochastic systems. Bayesian Anal. 4, 465-498

Jewell, C.P., Keeling, M.J. and Roberts, G.O. (2009b). Predicting undetected infections during the 2007 foot-and-mouth disease outbreak. J. R. Soc. Interface 6, 1145-1151

Jewell, C.P. and Roberts, G.O. (2012). Enhancing Bayesian risk prediction for epidemics using contact tracing. Biostatistics 13, 567-579

Kao, R.R. (2002). The role of mathematical modelling in the control of the 2001 FMD epidemic in the UK. Trends Microbiol. 10, 279-286

Keeling, M.J., Woolhouse, M.E.J., Shaw, D.J., Matthews, L., Chase-Topping, M., Haydon, D.T., Cornell, S.J., Kappey, J., Wilesmith, J. and Grenfell, B.T. (2001). Dynamics of the 2001 UK Foot and Mouth Epidemic: Stochastic Dispersal in a Heterogeneous Landscape. Science 294, 813-817

Minh, P., Stevenson, M., Jewell, C., French, N. and Schauer, B. (2011). Spatio-temporal analyses of highly pathogenic avian influenza H5N1 outbreaks in the Mekong River Delta, Vietnam, 2009. Spat. Spatiotemporal Epidemiol. 2, 49-57

National Audit Office (2002). The 2001 Outbreak of Foot and Mouth Disease. The Stationary Office, London

Ster, I.C. and Ferguson, N.M. (2007). Transmission parameters of the 2001 foot and mouth epidemic in Great Britain. PLoS One 2, e502

Tildesley, M.J., Deardon, R., Savill, N.J., Bessel, P.R., Brooks, S.P., Woolhouse, M.E.J. and Grenfell, B.T. (2008). Accuracy of models for the 2001 foot-and-mouth epidemic. Proc. Biol. Sci. 275, 1459-1468

WHAT, WHERE AND WEATHER? INTEGRATING OPEN-SOURCE TAXONOMIC, SPATIAL AND CLIMATOLOGIC INFORMATION INTO A COMPREHENSIVE DATABASE OF LIVESTOCK INFECTIONS

K.M. MCINTYRE[*], C. SETZKORN, M. WARDEH, P. J. HEPWORTH, A.D. RADFORD AND M. BAYLIS

SUMMARY

As 6 out of every 10 human pathogens came from animals, it seems likely that many future emerging threats to human health will also come from animals. A valuable open-source evidence-based database (the ENHanCEd Infectious Diseases database - EID2) has recently been created. The EID2 automatically extracts and analyses material contained in the metadata of millions of nucleotide sequences and in publications, and using this information we can draw conclusions that a certain type of pathogen infects a certain host, and is present in a certain country. Within this study, evidence from the EID2 of human pathogens and those of domestic animals was compared with that collated using non-automated techniques described previously. The results demonstrated that the information necessary for meta-analytical studies can be harnessed very efficiently using EID2 methodologies. Future work aims to expand the EID2, to build a more comprehensive list of livestock-associated and other pathogens.

INTRODUCTION

A great deal of time and resources have been put into studying individual pathogens or groups of pathogens affecting the animals with which humans have the most contact, because they potentially affect food security including socio-economic impacts and zoonotic disease transmission. For instance, it is estimated that one in four people in the UK suffer from diarrhoeal disease every year (Tam et al., 2012). As previous work has suggested that around 60% of infectious organisms known to be pathogenic to humans are zoonotic (Taylor et al., 2001; Woolhouse & Gowtage-Sequeria, 2005), it is likely that many of these episodes of illness are a result of erstwhile transmission from animals, causing emerging infections in humans.

Just over ten years ago, the first attempt to produce an inventory of the pathogens of humans was undertaken (Taylor et al., 2001). At the same time, this list was combined with further information for livestock, dogs, cats and wildlife, in a study which aimed to quantify

---

[*] K. Marie McIntyre, Department of Epidemiology and Population Health, Institute of Infection and Global Health, University of Liverpool, Leahurst Campus, Neston, Cheshire, CH64 7TE, UK. Email: k.m.mcintyre@liv.ac.uk

both pathogen characteristics and their interactions with key features of pathogen-host epidemiology including host range, zoonotic or emerging disease status and socio-economic importance (Cleaveland et al., 2001). Emerging diseases were defined as those that have appeared in a human or domestic mammal population for the first time, or have occurred previously but are increasing in incidence or expanding into areas where they had not previously been reported. The main material for both of those studies came from gathering specific information from textbooks. However, scientific literature was also examined to ascertain emerging infections. Such sources are potentially biased towards clinical infections and the data gathering would have been a lengthy process. There may be gaps in knowledge and therefore biases in animal and particularly companion animal disease data due to a lack of joined-up surveillance, which several recent projects aim to rectify (Moore et al., 2004a, b; Paiba et al., 2007; Radford et al., 2010). In addition, the job of identifying publications on pathogens and their hosts may be made more difficult due to our domestication of animals. For example, pathogens are found in domestic hosts in which they would not, naturally, be expected to occur, and husbandry practices can introduce host or pathogen vector species into new areas, change host susceptibility or behaviour, and expose hosts to new pathogens via modified transmission routes. Previous control programmes using, for instance antibiotics, can also promote further evolution of pathogens, changing their pathogenesis in host populations.

Within the present study, a recently created evidence-based resource has been utilised, the ENHanCEd Infectious Diseases database (EID2) (University of Liverpool, 2011), in order to provide a large, automated and systematically generated meta-analytical approach to studying the main pathogens and hosts involved in disease transmission within the domestic environment. The main way in which this study differs from that of Cleaveland et al. (2001) is in the use of the EID2 source for pathogen information. All evidence within the EID2 comes from, and is linked to, previously published sources; the database automatically extracts and analyses material contained in the metadata of millions of nucleotide sequences and in publications, storing the results in a hierarchical phylogenetic tree structure. The EID2 presents this information, and using it we can draw conclusions; for example, that a certain type of pathogen infects a certain host species, and is/was present in a certain country at a certain time. Therefore, where Cleaveland et al. (2001) used textbooks to find pathogens of specific hosts, in the EID2 approach the evidence comes from individual reports. The semi-automated nature of information gathering has also meant: a much larger quantity of proof of pathogens interacting with hosts has been used as evidence; it has been possible to study many more pathogens; and it has been possible to use a more exhaustive list of domestic animal hosts. In addition, automated methods to collate spatial information for pathogens assigned at the country level have been built into the EID2. Within this study, we have examined differences in the evidence for the occurrence of pathogens in Europe compared to the rest of the world.

This study aimed to (1) describe the methodologies used to create the EID2 database, and investigate the validity of using semi-automated literature gathering techniques to ascertain the countries within which pathogens occur, (2) identify the pathogens and hosts at the root of human and domestic animal disease networks to facilitate an understanding of the organisms to target for disease control; within conservation it is such pathogens which are of the greatest concern (Murray et al., 1999; Daszak et al., 2000; Cleaveland et al., 2001), and (3) examine the characteristics of these hosts and pathogens to distinguish potential drivers of disease transmission.

MATERIALS AND METHODS

The ENHanCEd Infectious Diseases (EID2) database

Individual host, vector and pathogen information: Identification of European pathogens was undertaken by creating an evidence-based database resource, the EID2 (University of Liverpool, 2011), in which evidence of the pathogens of all animals including humans (but not including fish) and their occurrence not just in Europe, but around the globe, was stored. The EID2 is a relational database which was designed via classes using domain driven paradigm. As of June 2012, it consists of 38 tables and 31 classes. Its core is the NCBI Taxonomy database (National Center for Biotechnology Information, 2012c) which provides a hierarchal structure based on the phylogenetic tree to information on each host, vector or pathogen (here-after referred to as ‘organism’) node, such that outputs can be obtained for organisms and higher taxonomic groups (for instance flaviviruses, ruminants). The information on each organism node includes evidence of alternative names or synonyms, which were mostly provided as a part of the information from the NCBI Taxonomy database. If phylogenetic information for an organism had not previously been included within the NCBI Taxonomy database structure, it could be added artificially into the EID2 at the correct place within the phylogenetic tree. Further information about each organism, such as its taxonomic rank (genus, species etc.) or taxonomic division for pathogens (i.e. bacteria – including rickettsia, fungi – including algal pathogens, helminths – including thorny-headed worms and pentastomids, protozoa, and viruses – including prion agents) has then been stored using a series of statements which are linked into the classes. These statements can be created using semi-automated methods such as data-mining of meta-data held within the NCBI Taxonomy database or the NCBI Nucleotide database (National Center for Biotechnology Information, 2012b), or they can be nominated by an individual using evidence from a publication. Data on publications described within PubMed (PubMed) are also held as a part of the EID2, with papers included based upon the organisms which they describe, and abstract information available for recently published papers. The EID2 has been set up in such a way that automated literature searches to look for specific topics associated with pathogens, for instance climate change, can be undertaken. In addition, climate data have been incorporated into the structure of the database, so that a number of climate variables are available to match with spatial data for pathogens, and the EID2 automatically models the climate conditions that determine pathogen distribution.

Pathogenic status of pathogens: In order to decide which pathogens cause significant clinical disease in hosts and therefore which need to be considered from a health and well-being perspective, a pathogenic status was manually assigned to each pathogen node included within the EID2. Definitions of pathogenic include:

- Frequently pathogenic – An organism which frequently has a clinically pathogenic effect (causes morbidity or mortality) upon humans or domestic animals.
- Non-pathogenic – An organism which causes no clinical signs within any of its hosts.
- Unknown pathogenicity – An organism for which there is insufficient evidence to decide whether it causes pathogenic effects in any host.

Information on pathogen-host and pathogen-location interactions: Specific information on pathogens affecting a certain host (termed a ‘host-pathogen interaction’) or pathogens occurring within a country (an ‘organism-country interaction’) was mined from meta-data held within the NCBI Nucleotide database (National Center for Biotechnology Information, 2012b). The last update from the nucleotide database was undertaken in December 2011. In addition, specific scientific publications were used as evidence of certain pathogen-host interactions.

A further source of information utilised for organism-country interactions came from automated searches of the PubMed database (PubMed) and the NCBI MeSH library (National Center for Biotechnology Information, 2012a); when the name of an organism and the (minor subject) MeSH term for a country co-occurred within a certain number of publications, an assumption was made about the occurrence of that organism within that country. Within the EID2, spatial data on organisms are hierarchically organised according to the NCBI MeSH library (National Center for Biotechnology Information, 2012a), thus allowing outputs at different regional levels. Spatial outputs can be in the form of a list or a map.

Humans and domestic animal species shortlist

Once the EID2 had been populated with information on organisms including pathogens, a short-list of humans, and domestic animal hosts with which we have close contact in Europe was drawn up. Data for this host population was examined for the purposes of this study. The list included domestic animals we eat or companion animals we keep as pets, and exotic animals also used as food sources or as pets (Table 1).

Table 1. Domestic animal host species including humans for which pathogens were studied

Scientific name	Common name	Scientific name	Common name
Agapornis personata	Masked lovebird	*Lama glama*	Lama
Agapornis roseicollis	Rosy-faced lovebird	*Lama pacos*	Alpaca
Anas platyrhynchos	Domestic duck	*Meleagris gallopavo*	Turkey
Anser anser	Domestic goose	*Melopsittacus undulatus*	Budgerigar
Bison bison	American bison	*Meriones unguiculatus*	Mongolian gerbil
Bison bonasus	European bison	*Mesocricetus auratus*	Syrian golden hamster
Bos indicus	Zebu	*Mus musculus*	House mouse
Bos taurus	Cow	*Mustela putorius furo*	Domestic ferret
Camelus dromedarius	Dromedary	*Numida meleagris*	Helmeted guineafowl
Canis lupus familiaris	Domestic dog	*Nymphicus hollandicus*	Cockatiel
Capra hircus	Domestic goat	*Oryctolagus cuniculus*	Domestic rabbit
Capreolus capreolus	Roe deer	*Ovis aries*	Sheep
Cavia porcellus	Domestic guinea pig	*Ovis aries musimon*	Mouflon
Cervus elaphus	Red deer	*Pavo cristatus*	Blue peafowl
Chinchilla lanigera	Chinchilla	*Phasianus colchicus*	Ring-necked pheasant
Columba livia	Domestic pigeon	*Rangifer tarandus*	Reindeer
Cricetus cricetus	Common hamster	*Rattus norvegicus*	Brown rat
Dama dama	Fallow deer	*Rattus rattus*	Black rat
Equus asinus	Domestic donkey	*Rhombomys opimus*	Great Gerbil
Equus caballus	Domestic horse	*Serinus canaria*	Canary
Felis catus	Domestic cat	*Struthio camelus*	Ostrich
Gallus gallus	Chicken	*Sus scrofa*	Wild boar
Homo sapiens	Humans	*Sus scrofa domesticus*	Domestic pig
Lagopus lagopus scotica	Red grouse		

Statistical analyses of data

The validity of using automatically mined pathogen-location information in the EID2 database: The threshold number of papers in which to infer a pathogen-location interaction from automated searches of the PubMed database (PubMed) and the NCBI MeSH library (National Center for Biotechnology Information, 2012a) was investigated using two different approaches.

First, the positive predictive value (PPV) of a test for whether papers showed a 'true' relationship between a pathogen and country MeSH term was established by manually analysing papers. This involved using a sub-sample of papers stratified according to the pathogen and continent on which the MeSH term country was located, after which papers were selected for inclusion in PPV calculation using a random number generator. Differences in the likelihood of a truly positive assumed relationship (not previously described using NCBI Nucleotide database metadata (National Center for Biotechnology Information, 2012b)) arising as a result of pathogenic status or taxonomic division were examined using generalised linear models (GLM) with binomial errors and logit link functions.

Second, a GLM with binomial errors and a logit link function was used to ascertain the likelihood of an assumed pathogen-country interaction being correct. The outcome variable within the model was if a pathogen-country interaction had been reported in metadata from the NCBI Nucleotide database for at least one nucleotide sequence. The explanatory variable was the number of papers from a PubMed search in which a pathogen and country MeSH term had co-occurred. Non-linear relationships in the data were investigated using generalised additive modelling (GAM) and inclusion of polynomial terms or a linear spline function. The break points within the spline function were explored using an iterative process where the break increased from the minimum in discrete steps representing the number of papers. Significantly improved GLM model fits were established by comparing models using Chi-squared analysis; the final GLM model was ascertained using deviance residuals with the smallest value being the best fit. Further to the linear spline model, a model in which the number of papers from PubMed was recoded as a factor was also tested, again exploring breaks using an iterative process. This technique would allow simple interpretation of the characteristics of each part of the relationship between nucleotide sequences and PubMed papers.

Pathogen range within hosts: GLMs with binomial errors and logit link functions were used to explore if the likelihood of a human pathogen being emerging was influenced by: the number of hosts it occurred in (one host, two hosts or more than two hosts) and whether the pathogen was zoonotic or non-zoonotic. Information on the taxonomic division of pathogens and their pathogenicity was also included within the analyses as covariates. Emerging and zoonotic status was taken from Taylor et al. (2001) and Woolhouse & Gowtage-Sequeria (2005). The models were built using stepwise deletion of terms by comparing models using Chi-squared analysis, and statistical significance was determined by a *P*-value of less than 0.05. Adjusted odds ratios (AOR) significantly different from one were used as an indicator of raised or lowered likelihood of pathogens being emerging. Hosmer-Lemeshow goodness-of-fit tests were used to judge the goodness-of-fit of the models. The influence of number of hosts and emerging status upon the likelihood of pathogens being zoonotic compared to non-zoonotic was explored using models which included the same covariates and methodologies as described above for emerging pathogens.

Differences in spatial representation of pathogen species: The influence of number of hosts, emerging and zoonotic status upon the likelihood of pathogens being found in Europe compared to the rest of the world was explored using models which used the same methodologies as described above for emerging pathogens, and including a taxonomic division covariate.

RESULTS

The validity of using automatically mined pathogen-location information in the EID2 database

Positive predictive value approach: If there was no direct evidence available from the NCBI Nucleotide database, a threshold (*t*) of number of papers from PubMed in which a pathogen name and MeSH term for a country co-occurred within each paper was used as evidence of pathogen prevalence. This was based upon a preliminary study in which papers had been stratified according to the pathogen and the continent to which they were linked via a MeSH term for a country. The evidence within papers was then manually checked to substantiate that the pathogen was found in hosts within a MeSH term country. This allowed the calculation of the positive predictive value (PPV) of the test (1-((1-PPV)^*t*)). The PPV for each paper was 0.95 (SE=0.05), and so a threshold of five papers was chosen which would equate to 99.9% certainty of a ‘true’ relationship. The likelihood of an assumed relationship between a pathogen and a MeSH term country being true was not affected by pathogenic status (P=0.393) or taxonomic division (P>0.05).

Binomial regression modelling approach: Having explored GAM results and a fifth order polynomial term within GLMs, a final binomial regression model included a $\log_{10}(n+1)$ transformation of the explanatory variable and a linear spline function with two breaks. The regression results suggest that the number of papers identified using PubMed is positively related to the likelihood of a pathogen-country interaction having been reported in the NCBI Nucleotide database, but the significance and characteristics of this relationship change dependent upon the number of PubMed papers (Fig. 1: Overall results, df=3, 3225, left-hand section of the line P<0.001, middle section P<0.001, right-hand section P=0.132). The alternative model in which the number of papers from PubMed was recoded as a factor explained less of the residual deviance in the model (3989.1 as opposed to 3974.8, respectively); however, it too suggested two breaks points. Within this model, the odds of at least one nucleotide sequence describing a pathogen-country interaction increased significantly (P<0.001) by 1.86 times (Confidence Interval (CI): 1.55-2.22) when there were between two and 12 PubMed papers in which a pathogen name and MeSH term for a country co-occurred, compared to when the interaction was described by only one PubMed paper. The odds of at least one nucleotide sequence increased significantly (P<0.001) to 2.89 times (CI: 2.35-3.55) compared to one PubMed paper when the number of PubMed papers was more than 12.

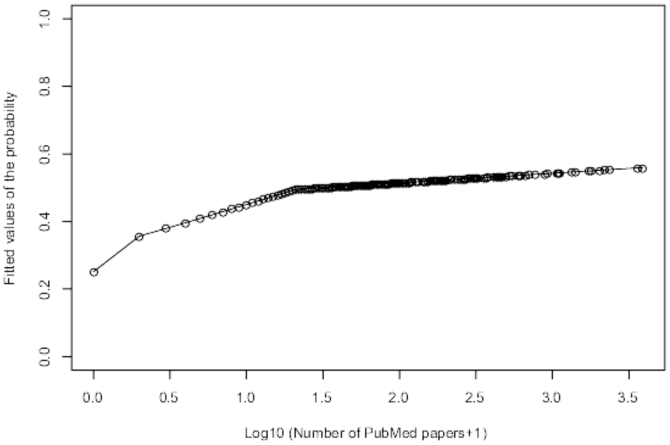


Fig. 1 The relationship between the number of papers identified using PubMed and the probability of a pathogen-country interaction having been reported in the NCBI Nucleotide database. The final generalised linear regression model with binomial errors and logit link function included a $\log_{10}(n+1)$ transformation of the explanatory variable (number of PubMed papers), and a linear spline function fitted with two breaks

Summary of the EID2 database

Pathogen range within hosts: Of the final short-list of domestic animal hosts with which we have close contact in Europe including humans, the EID2 described interactions connecting 47 hosts to 2597 pathogen species using meta-data held within the NCBI Nucleotide database. Within the host species, 793 (30.5%) pathogens were bacteria species, 395 (15.2%) fungi, 705 (27.1%) helminths, 372 (14.3%) protozoa and 332 (12.8%) viruses. Pathogens generally affected very few different hosts, with the frequency of pathogen species to hosts characterised by a negative binomial distribution (median value=1, mean value=1.63, max value=21). Most (70.9%) of the pathogen species interacted with one host, 86.0% affected up to 2 hosts, and 98.6% affected six or fewer. The greatest proportion of pathogens (49.8%) affected humans only, 32.5% affected domestic animals and 17.8% affected both humans and animals. Of the human pathogens, 74.9% affected humans only and 25.1% affected humans and animals. Of the animal pathogens, 65.7% affected animals only and 34.3% affected humans and animals. The pathogens affecting the greatest number of hosts are presented in Table 2, including all those with at least seven hosts.

Comparing pathogens which are emerging (n=169) with those not emerging (n=1187) (Table 3a) indicated that the odds of emergence differed by taxonomic division: helminth pathogens were found to be 80% less likely, protozoans twice more likely and viruses six times more likely to emerge, than bacteria. Furthermore, the odds of pathogen emergence were five times higher if they had more than two compared to one host species associated with them, and emerging pathogens were 64% more likely to be zoonotic than non-zoonotic. Pathogenicity did not significantly affect the odds of a pathogen being emerging.

Comparing pathogens which are zoonotic (n=829) with those non-zoonotic (n=527) (Table 3b) indicated that the odds of a pathogen being zoonotic differed by taxonomic division: helminth pathogens were 16 times, protozoans twice and viruses nearly four times more likely to be zoonotic than bacteria. Moreover, the odds of pathogens being zoonotic were three and a half times higher if they had two, and nearly 11 times higher if they had more than two hosts compared to one host species associated with them. Pathogenicity did not significantly affect the odds of a pathogen being zoonotic. Emerging status caused the model to become unstable and it was therefore not included within the final results; when forced in, it increased the odds of pathogens being zoonotic.

Table 2. The number of human or animal hosts affected by pathogen species, including taxonomic division. *E* or *NE* after pathogen names denote emerging or not emerging and *Z* or *NZ* denote zoonotic or non-zoonotic status according to Taylor et al. (2001) and Woolhouse & Gowtage-Sequeria (2005). NA denotes pathogens not included in these earlier studies

Pathogen name	Pathogen type	Number of hosts	Pathogen name	Pathogen type	Number of hosts
Escherichia colli $_{E,Z}$	Bacteria	21	*Gongylonema pulchrum* $_{NE,Z}$	Helminth	9
Giardia intestinalis $_{E,Z}$	Protozoa	20	*Leptospira interrogans* $_{E,Z}$	Bacteria	9
Toxoplasma gondii $_{E,Z}$	Protozoa	18	Ovine Herpesvirus 2 $_{NA}$	Virus	9
Anaplasma phagocytophilum $_{E,Z}$	Bacteria	15	Rotavirus A $_{E,Z}$	Virus	9
Cryptosporidium parvum $_{E,Z}$	Protozoa	14	*Clostridium perfringens* $_{NE, Z}$	Bacteria	8
Rabies virus $_{E,Z}$	Virus	13	Cowpox virus $_{NE, Z}$	Virus	8
Staphylococcus aureus $_{E,Z}$	Bacteria	13	*Enterococcus faecalis* $_{E,Z}$	Bacteria	8
Neospora caninum $_{NA}$	Protozoa	12	*Enterococcus faecium* $_{E,Z}$	Bacteria	8
Echinococcus granulosus $_{E,Z}$	Helminth	11	*Enterocytozoon bieneusi* $_{E, NZ}$	Fungi	8
Borna Disease virus $_{NE,Z}$	Virus	10	Hepatitis E virus $_{E,Z}$	Virus	8
Newcastle Disease virus $_{NE,Z}$	Virus	10	*Malassezia sympodialis* $_{NE, Z}$	Fungi	8
Pasteurella multocida $_{NE,Z}$	Bacteria	10	*Brachyspira pilosicoli* $_{NA}$	Bacteria	7
Trypanosoma cruzi $_{E,Z}$	Protozoa	10	Influenza A virus $_{E, Z}$	Virus	7
Babesia divergens $_{NE,Z}$	Protozoa	9	*Mecistocirrus digitatus* $_{NE, Z}$	Helminth	7
Chlamydophila psittaci $_{NE,Z}$	Bacteria	9	*Pneumocystis carinii* $_{E, Z}$	Fungi	7
Cryptosporidium muris $_{NA}$	Protozoa	9	*Saccharomyces cerevisiae*	Fungi	7
Echinococcus canadensis $_{NA}$	Helminth	9	*Trichostrongylus colubriformis* $_{NE, Z}$	Helminth	7
Encephalitozoon cuniculi $_{E,Z}$	Fungi	9	West Nile virus $_{E, Z}$	Virus	7
Fasciola hepatica $_{NE, Z}$	Helminth	9			

Host range: The pathogens of humans and domestic animal hosts are characterised by 4,223 host-pathogen interactions, with the greatest number found in (decreasing order): humans, sheep/goats, cattle, small mammals, pigs, dogs, equids, cats, chickens, other edible birds, deer, exotic mammals, ducks and caged birds (Fig. 2). The breakdown of pathogen species by taxonomic division is given in Fig. 3. Pathogens of deer were most likely to cause significant clinical effects in one of their hosts as opposed to being non-pathogenic in any host, followed by those of (descending): dogs, cattle, pigs, humans, cats, ducks, exotic mammals, caged birds, equids, sheep/goats, other edible birds, small mammals and chickens.

Table 3. Logistic regression models for the likelihood of pathogens being (a) emerging or (b) zoonotic. Adjusted odds ratios (AOR) shown in bold differ significantly (*P*<0.05) from one

(a)

	AOR	95% Confidence Interval	
Number of hosts species			
1	baseline	-	-
2	1.41	0.75	2.63
>2	**4.82**	3.00	7.74
Taxonomic division			
Bacteria	baseline	-	-
Fungi	1.02	0.59	1.75
Helminths	**0.18**	0.09	0.37
Protozoa	**2.06**	1.00	4.21
Viruses	**6.16**	3.94	9.63
Zoonotic status			
Non-zoonotic	baseline	-	-
Zoonotic	**1.64**	1.07	2.53
Hosmer-Lemeshow goodness-of-fit test		*P*=0.47	

(b)

	AOR	95% Confidence Interval	
Number of hosts species			
1	baseline	-	-
2	**3.53**	2.18	5.72
>2	**10.92**	5.89	20.23
Taxonomic division			
Bacteria	baseline	-	-
Fungi	0.75	0.55	1.02
Helminths	**16.35**	9.04	29.59
Protozoa	**2.29**	1.17	4.50
Viruses	**3.77**	2.56	5.56
Hosmer-Lemeshow goodness-of-fit test		*P*=0.54	

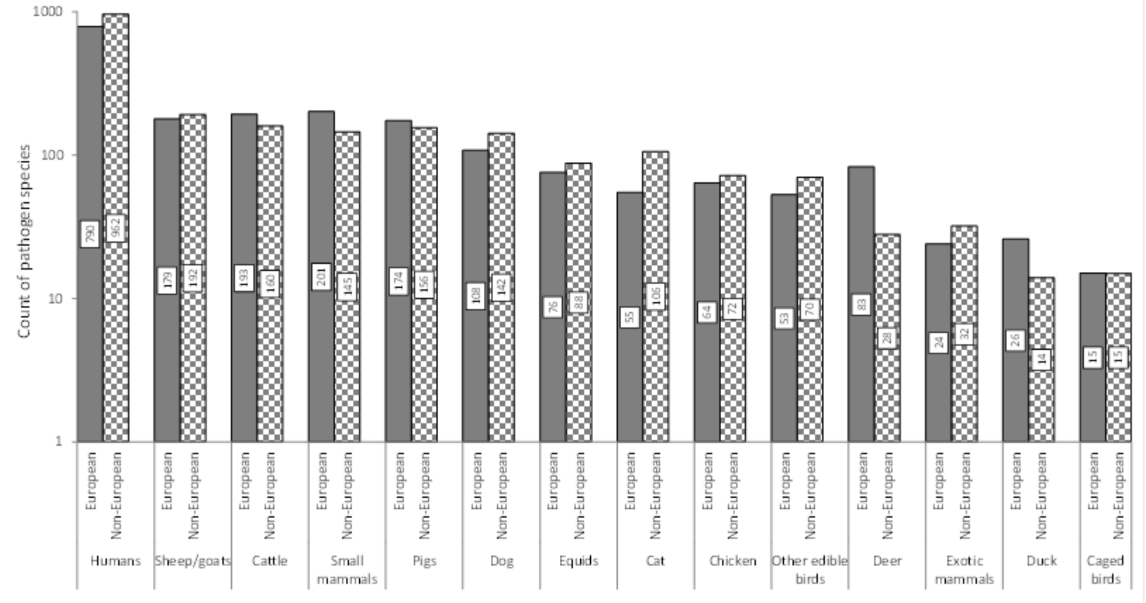
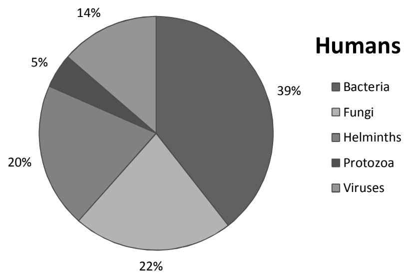
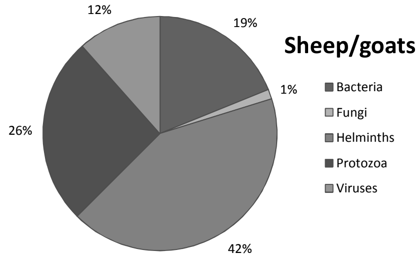


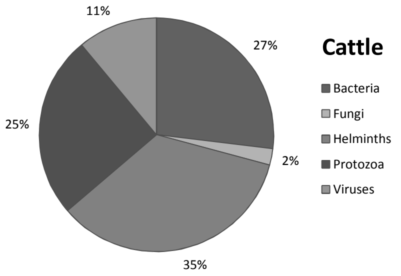
Fig. 2 Count of human or animal host and pathogen species interactions (summarised by host groups) in EID2, including whether pathogens occurred in Europe or the rest of the world. Data labels are for the number of pathogen species

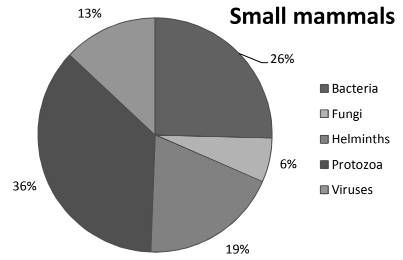
<u>Differences in spatial representation of pathogen species:</u> The number of pathogen species recorded varied by European country (Fig. 4). In all host groups except cattle, small mammals, pigs, deer and ducks, the number of pathogens reported was greater in the rest of the world than in Europe (Fig. 2). In all host groups except humans and cattle, the proportion of frequently pathogenic compared to non-pathogenic pathogens was higher in Europe than in the rest of the world.

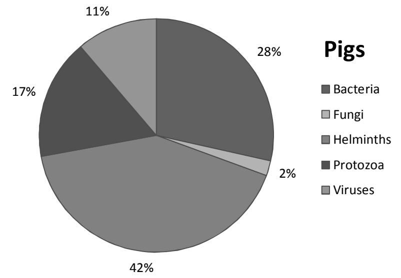
Comparing pathogens in Europe (n=601) as opposed to in the rest of the world (n=755) (Table 4) indicated that the odds of a pathogen being found in Europe differed by taxonomic division: helminth pathogens were found to be 80%, protozoans half and viruses two and a half times less likely to be found in Europe compared to the rest of the world than bacteria. Furthermore, the odds of pathogens being found in Europe were three times or six and a half times higher if they had two or more than two hosts compared to one host species associated with them, respectively. Emerging pathogens were more than three and a half times more likely to be found in Europe than the rest of the world than pathogens not emerging. Zoonotic status did not significantly affect the odds of a pathogen being found in Europe.

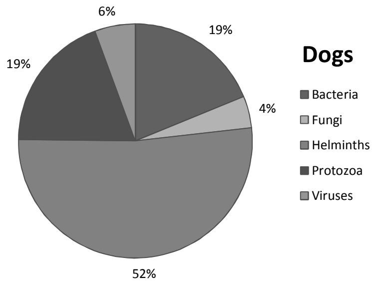

Humans
14%
5%
20%
22%
39%
Bacteria
Fungi
Helminths
Protozoa
Viruses

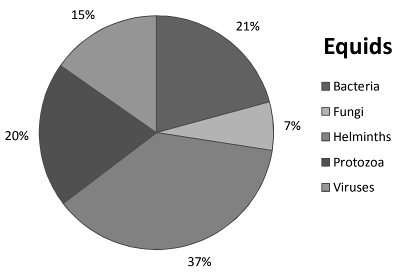

Sheep/goats
12%
19%
1%
26%
42%
Bacteria
Fungi
Helminths
Protozoa
Viruses

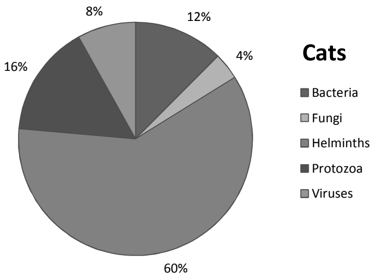

Cattle
11%
27%
25%
2%
35%
Bacteria
Fungi
Helminths
Protozoa
Viruses

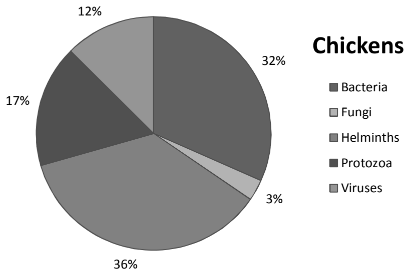

Small mammals
13%
26%
6%
36%
19%
Bacteria
Fungi
Helminths
Protozoa
Viruses

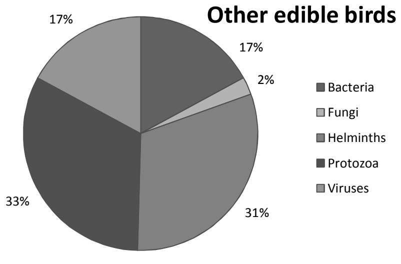

Pigs
11%
28%
17%
2%
42%
Bacteria
Fungi
Helminths
Protozoa
Viruses


Dogs
6%
19%
19%
4%
52%
Bacteria
Fungi
Helminths
Protozoa
Viruses


Equids
15%
21%
20%
7%
37%
Bacteria
Fungi
Helminths
Protozoa
Viruses


Cats
8%
12%
4%
16%
60%
Bacteria
Fungi
Helminths
Protozoa
Viruses


Chickens
12%
32%
17%
3%
36%
Bacteria
Fungi
Helminths
Protozoa
Viruses


Other edible birds
17%
17%
2%
33%
31%
Bacteria
Fungi
Helminths
Protozoa
Viruses

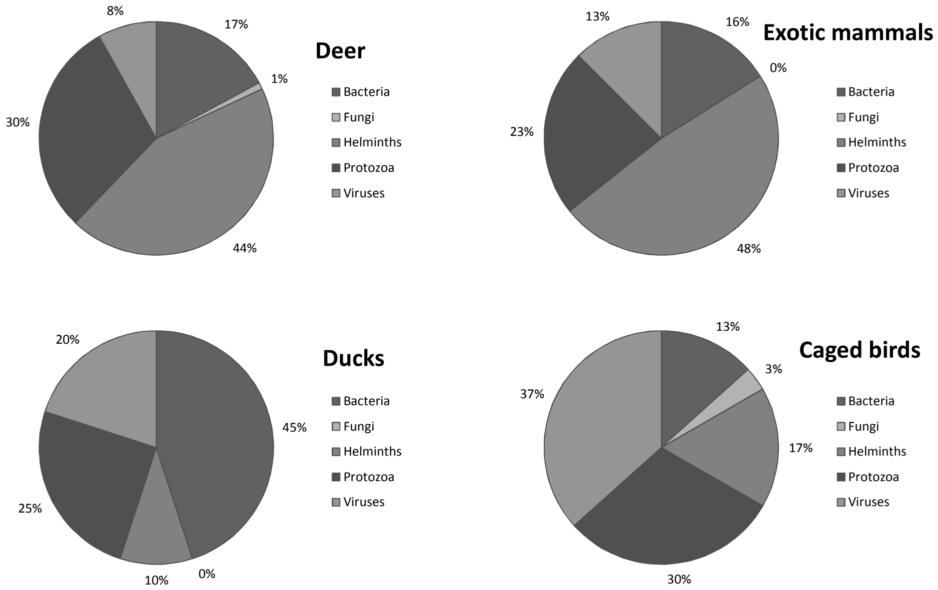


Fig. 3 Proportion of pathogen species by taxonomic division for different host groups

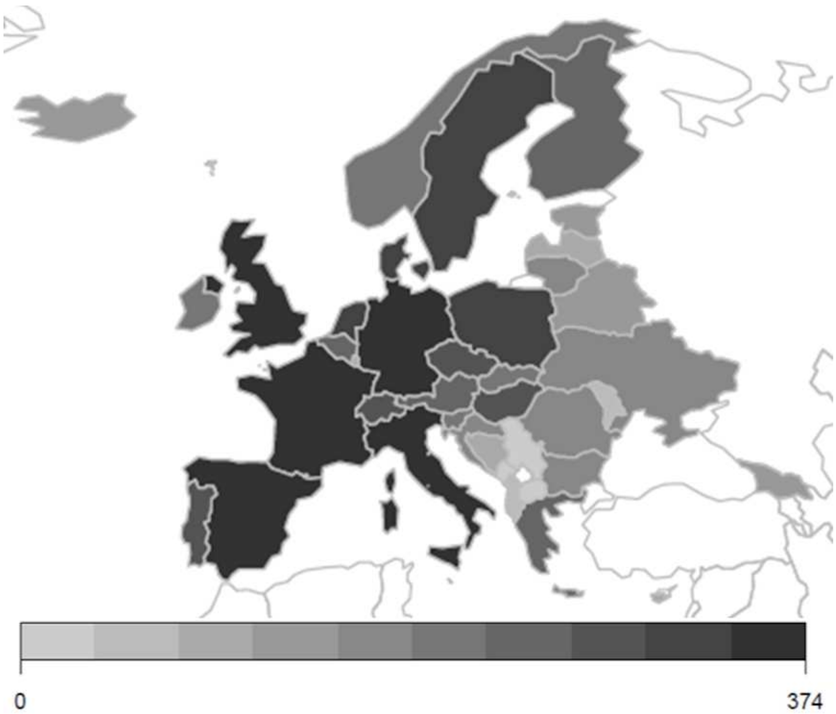


Fig. 4 Map of Europe showing number of pathogen species in each country. Scale depicts number of species

Table 4. Logistic regression model for the likelihood of pathogens being found in Europe as opposed to the rest of the world. Adjusted odds ratios (AOR) shown in bold differ significantly (*P*<0.05) from one

	AOR	95% Confidence Interval	
Number of hosts species			
1	baseline	-	-
2	**2.97**	1.97	4.46
>2	**6.66**	4.44	10.01
Taxonomic division			
Bacteria	baseline	-	-
Fungi	0.94	0.70	1.26
Helminths	**0.17**	0.11	0.25
Protozoa	**0.48**	0.26	0.91
Viruses	**0.39**	0.27	0.59
Emerging status			
Not emerging	baseline	-	-
Emerging	**3.68**	2.43	5.58
Hosmer-Lemeshow goodness-of-fit test		*P*=0.22	

DISCUSSION

The EID2 database described utilises the individual reports of host-pathogen interactions provided by genome, gene and transcript sequence recording resources, in this case the NCBI Nucleotide database (National Center for Biotechnology Information, 2012b). This study demonstrates the vast amount of hitherto untapped information which can be used for meta-analytical approaches to examine the characteristics and drivers of pathogens, and particularly those affecting multiple hosts. It illustrates that by using such methods we can examine differences in surveillance of pathogens but also begin to untangle the vast networks of hosts in which they occur. There are biases inherent in the use of sequence databank resources, however, including likely under-reporting of pathogens for various reasons: they might not cause clinical infection and thus reporting would not have tangible benefits, they might not cause a notifiable disease or their geographical distribution or host range might have changed so that they would not be sought in host species; they might be prevalent in countries in which surveillance resources are restricted meaning a lack of identification but also submissions to sequence databases would not occur; and for some pathogen groups such as fungi, it might be difficult or unnecessary in clinical practice to identify species, when a cure can be secured without extra effort. Another bias might be in the hosts in which pathogens are reported: it is likely that some species such as humans will be relatively over-represented whereas little may be known or reported on the pathogens of wildlife; and pathogens may be reported from research on laboratory animals rather than after infection of natural hosts. Furthermore, biological material for certain pathogens may be easier to obtain than for others, either because they are physically difficult or costly to isolate within a host.

The validity of using automatically mined pathogen-location information in the EID2 database

The initial analyses within this study validate the use of semi-automatic data-mining processes within the EID2 to ascertain where pathogens occur at the country level, given consideration of sample sizes. Future developments of the database will include undertaking similar tests for automatic assignation of host-pathogen interactions, and increasing the spatial resolution of pathogen-location information.

Pathogen range within hosts

Considering the proportion of pathogens found within certain taxonomic divisions, the results for bacteria, fungi and helminths were largely similar for this study in comparison to previous work examining human, livestock and domestic carnivore populations (Cleaveland et al., 2001). The proportions of protozoan species were higher, however, and viruses were slightly lower, possibly due to the additional inclusion of small mammal, domestic bird and exotic (domestic animal) host species. This study also examined pathogens not causing clinical infection or with unknown aetiologies, where the other analysis used only clinical disease.

Within the study, the proportion of pathogens affecting more than one host was much lower than for Cleaveland et al. (2001) (29.1% compared to 62.7%), perhaps due to the inclusion of non-clinical infection but also because nucleotide sequences are more likely to be sent into a sequence databank if they have been found for the first time in a novel host or unusual environment, potentially skewing the distribution of data on pathogens affecting host species. Cleaveland et al. (2001) reported 39.1% of human pathogens infecting domestic animal hosts; a higher proportion than found within this study (25.1%) because given the relative investment in healthcare systems, information on non-clinical disease is more likely to be ascertained for human than animal hosts. This would further explain the slightly lower proportion of livestock pathogens affecting humans in this than the Cleaveland study (34.3% and 39.4%, respectively). That said, a conclusion that fewer multi-host pathogens cause disease in humans than domestic animals was found in both studies, and postulated upon by Cleaveland et al. (2001). For the purposes of presentation, some of the most frequently recorded zoonotic pathogens in Europe were not included within the list of those associated with the greatest number of hosts, for instance *Listeria monocytogenes* (connected to six hosts), *Campylobacter coli*, *C. jejuni* and *Salmonella enterica* (all connected to five hosts). A more complete version of this table can be requested from the authors.

The results of comparison of emerging and not emerging pathogens have corroborated previous work: pathogen emergence is associated with number of host species, with the likelihood of emergence increasing with greater numbers (Cleaveland et al., 2001; Woolhouse & Gowtage-Sequeria, 2005); helminths were less likely to emerge and viruses and protozoa more likely (Cleaveland et al., 2001; Dobson & Foufopoulos, 2001; Taylor et al., 2001), though a reduced risk of emergence for fungi (Cleaveland et al., 2001; Taylor et al., 2001) was not observed in this study, perhaps as risks were relative to bacteria; and the odds of pathogens being emerging compared to not emerging were higher for zoonotic as opposed to non-zoonotic pathogen species (Cleaveland et al., 2001; Taylor et al., 2001; Woolhouse & Gowtage-Sequeria, 2005; Jones et al., 2008).

That zoonotic compared to non-zoonotic pathogens were influenced by number of host species reflects the capacity of a pathogen to spread or transmit within host populations; disease transmission is more likely given a network of hosts which may be closely phylogenetically related or which interact closely as with domesticated animal populations (Karesh et al., 2012).

Host range and the spatial representation of pathogen species

Many of the results in this study for host range of pathogens and the countries in which each pathogen species occurred may be influenced by sampling biases in surveillance. For instance, the number of host-pathogen interactions identified within host groups reflects differences in healthcare, politics and the funding of disease surveillance (Daszak et al., 2000). Patterns in the proportion of clinical effects within host groups likewise could be due to biases in non-clinical pathogens not being isolated within certain host species, for instance deer, or being less likely to be isolated in, for example, developing world countries compared to Europe. Spatial sampling biases are reflected in the map illustration in which eastern European countries had much fewer numbers of pathogen species compared to western European; pathogen isolations from Eastern Europe before relatively recent country boundary changes may not have been included in our data. Furthermore, comparison of pathogens in Europe and the rest of the world suggests differences in sampling effort between taxonomic divisions, when fewer host species were associated with pathogens, and when pathogen species had been awarded a politically important status: they were emerging. This final result might be surprising given the relatively recent plethora of publications and funding on emerging pathogens (Taylor et al., 2001; Woolhouse & Gowtage-Sequeria, 2005; Murphy, 2008; Woolhouse, 2008), much of which have come from North America (Daszak et al., 2000, 2001; Dobson & Foufopoulos, 2001; Jones et al., 2008). It might also, however, reflect the sheer quantity of pathogen isolations undertaken in areas outside Europe.

CONCLUSIONS

Previous meta-analytical studies and reviews suggest that in order to provide early warning systems for disease emergence and to study zoonotic pathogens, a united effort is needed to collate information on their complex epidemiologies (Daszak et al., 2000; Cleaveland et al., 2001; Taylor et al., 2001; Woolhouse & Gowtage-Sequeria, 2005; Wolfe et al., 2007; Murphy, 2008; Karesh et al., 2012). It is hoped that the work described within this study goes some way to demonstrating that much of the necessary information is potentially already available and can be harnessed using semi-automated methodologies such as that provided by the EID2; by being more methodical within our unified efforts, we can build ‘one health’ disease surveillance systems.

Future work aims to expand the information held within the EID2, to build a more comprehensive list of livestock and other pathogens. Specific improvements will include: greater spatially detailed information for pathogens; improvements to the database’s ability to handle records with badly defined host species; the addition of further environmental data to produce better models to explain pathogen distributions and predict them in the future, given climate change; and information to allow users to work at the level of diseases, rather than individual pathogens or groups of pathogens. Potentially, the EID2 will benefit the veterinary and human health communities by providing greater lead-time for pathogen surveillance or for control measure design, and it will help inform clinicians about pathogens driving clinical disease; their origins, a temporal and spatial indication of recent disease outbreaks and links to publications which describe them.

ACKNOWLEDGEMENTS

The authors would like to thank Helen Roberts (UK Department for Environment, Food and Rural Affairs) and John Stephenson (Health Protection Agency) for their input during the development of the EID2 database, and Sally Eagle for help with the piecewise regression modelling. This work was funded by NERC through the ERA-ENVHEALTH network and by a BBSRC Strategic Tools and Resources Development Fund grant.

REFERENCES

Cleaveland, S., Laurenson, M.K. and Taylor, L.H. (2001). Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence. Phil. Trans. R. Soc. B 356, 991-999

Daszak, P., Cunningham, A.A. and Hyatt, A.D. (2000). Wildlife ecology - Emerging infectious diseases of wildlife - Threats to biodiversity and human health. Science 287, 443-449

Daszak, P., Cunningham, A.A. and Hyatt, A.D. (2001). Anthropogenic environmental change and the emergence of infectious diseases in wildlife. Acta Trop. 78, 103-116

Dobson, A. and Foufopoulos, J. (2001). Emerging infectious pathogens of wildlife. Phil. Trans. R. Soc. B 356, 1001-1012

Jones, K.E., Patel, N.G., Levy, M.A., Storeygard, A., Balk, D., Gittleman, J.L. and Daszak, P. (2008). Global trends in emerging infectious diseases. Nature 451, U990-U994

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

Moore, G.E., Ward, M.P., Dhariwal, J., Wu, C.C., Glickman, N.W., Lewis, H.B. and Glickman, L.T. (2004a). Development of a national companion animal syndromic surveillance system for bioterrorism. Gisvet '04: Second International Conference on the Application of GIS and Spatial Analysis to Veterinary Science, 9-11

Moore, G.E., Ward, M.P., Dhariwal, J., Wu, C.C., Glickman, N.W., Lewis, H.B. and Glickman, L.T. (2004b). Use of a primary care veterinary medical database for surveillance of syndromes and diseases in dogs and cats. J. Vet. Intern. Med. 18, 386-386

Murphy, F.A. (2008). Emerging zoonoses: The challenge for public health and biodefense. Prev. Vet. Med. 86, 216-223

Murray, D.L., Kapke, C.A., Evermann, J.F. and Fuller, T.K. (1999). Infectious disease and the conservation of free-ranging large carnivores. Anim. Conserv. 2, 241-254

National Center for Biotechnology Information, US National Library of Medicine (2012a). The NCBI Medical Subject Headings (MeSH) database homepage. http://www.ncbi.nlm.nih.gov/mesh

National Center for Biotechnology Information, US National Library of Medicine (2012b). The NCBI Nucleotide database homepage. http://www.ncbi.nlm.nih.gov/nuccore

National Center for Biotechnology Information, US National Library of Medicine (2012c). The NCBI Taxonomy database homepage. http://www.ncbi.nlm.nih.gov/Taxonomy/

Paiba, G.A., Roberts, S.R., Houston, C.W., Williams, E.C., Smith, L.H., Gibbens, J.C., Holdship, S. and Lysons, R. (2007). UK surveillance: Provision of quality assured information from combined datasets. Prev. Vet. Med. 81, 117-134

PubMed National Center for Biotechnology Information (NCBI), National Library of Medicine (NLM), Bethesda, Maryland, US. http://www.ncbi.nlm.nih.gov/pubmed/

Radford, A., Tierney, A., Coyne, K.P., Gaskell, R.M., Noble, P.J., Dawson, S., Setzkorn, C., Jones, P.H., Buchan, I.E., Newton, J.R. and Bryan, J.G.E. (2010). Developing a network for small animal disease surveillance. Vet. Rec. 167, 472-474

Tam, C.C., Rodrigues, L.C., Viviani, L., Dodds, J.P., Evans, M.R., Hunter, P.R., Gray, J.J., Letley, L.H., Rait, G., Tompkins, D.S., O'Brien, S.J. and I. I.D. Study Executive Comm. (2012). Longitudinal study of infectious intestinal disease in the UK (IID2 study): incidence in the community and presenting to general practice. Gut 61, 69-77

Taylor, L.H., Latham, S.M. and Woolhouse, M.E.J. (2001). Risk factors for human disease emergence. Phil. Trans. R. Soc. B 356, 983-989

University of Liverpool (2011). The ENHanCEd Infectious Diseases database (EID2). National Centre for Zoonosis Research, www.zoonosis.ac.uk/eid2

Wolfe, N.D., Dunavan, C.P. and Diamond, J. (2007). Origins of major human infectious diseases. Nature 447, 279-283

Woolhouse, M.E.J. (2008). Epidemiology - Emerging diseases go global. Nature 451, 898-899

Woolhouse, M.E.J. and Gowtage-Sequeria, S. (2005). Host range and emerging and reemerging pathogens. Emerg. Infect. Dis. 11, 1842-1847

THE IMPACT OF Q FEVER CONTROL MEASURES: A QUANTITATIVE VALUE CHAIN ANALYSIS OF DUTCH DAIRY GOAT PRODUCTS

M.A. GONGGRIJP[*], M. NIELEN, J. RUSHTON AND H. HOGEVEEN

SUMMARY

Measures taken by a government to control an animal disease can have great impact on the actors of the relevant livestock industry. A value chain analysis, in combination with a gross margin analysis, was conducted to study the socio-economic impact of the control measures against Q fever in the Dutch dairy goat industry in December 2009. The value chain described dairy goat products from production to consumption complemented with amounts of products and money flows. The results demonstrated a major negative impact on gross margins of farmers and most of the actors directly related to them.

Value chain analyses often remain qualitative and descriptive. In this study, however, actors were willing to reveal quantitative information. Therefore it was possible to extend the value chain analysis with a quantification of the impact. It highlights that compensation of farmers alone may not adequately address the economic impact of animal diseases.

INTRODUCTION

During the last decades, livestock industries all over the world have undergone rapid developments, resulting in larger and more complex systems (Rushton, 2010). These livestock industries include many different actors involved in the production of a certain commodity from production level through several steps of processing to the consumption of the product. These actors are all related to one or more other actors in this industry, resulting in a dynamic network of services, transport, regulations, commodity- and money flows.

An animal disease outbreak can have a big economic impact on a livestock industry because of the loss of production due to sick and affected animals. Therefore governments often act on these events and try to suppress the disease with different control programmes. These measures will have an effect on the spread of the disease, but can also result in a negative economic impact on the industry. It is of great importance to know how the control programmes will affect a livestock industry from an economic and epidemiologic point of

---

[*] Maaike Gonggrijp, GD Animal Health Service, PO Box 9, 7400 AA Deventer, The Netherlands. Email: m.gonggrijp@gddeventer.com

view, because they will influence the behaviour and incentives of the producers and as a consequence, the economic situation (Rich, 2007). All the concerned actors in an industry will react differently to the impact of the control measures, but there will be differences between industries. The more resilient livestock industries will absorb the effects while other industries could become less efficient or collapse completely (FAO, 2011).

An example of an animal disease outbreak in which control measures had a great influence on a livestock industry, is the outbreak of Q fever in the Netherlands. This disease outbreak and the following control programmes had a dramatic effect on the Dutch dairy goat industry. This particular industry is relatively new in the Netherlands. Only for the last 25 years have goats been kept professionally in the Netherlands. This industry has grown enormously from 98,000 dairy goats in 2000 to 231,000 dairy goats in 2009 (information source: the Dutch Dairy Board). Most of the produced goat milk is exported as bulk milk or dairy products to other European countries or Asia. Only a small part is consumed as dairy products in the Netherlands. These dairy products consist of hard and soft cheese, yoghurt, milk and milk powder.

In 2005, *Coxiella burnetii*, the causal zoonotic bacterium of Q fever, was diagnosed for the first time in the Netherlands as a cause of abortion problems on a dairy goat farm. In the following years, more and more goat herds got infected and in 2007 the first Q fever outbreak in humans was diagnosed. Since then thousands of people got infected with the bacterium, with the outbreak peaking in 2009. To control the outbreak by lowering the infection rate and reducing the excretion of bacteria in the environment, (pregnant) goats from infected farms were euthanised and a vaccination programme was started. Also, animal movement restrictions, a breeding ban, bulk milk monitoring and extra hygiene programmes became mandatory. These combined control strategies seemed to have a positive effect on the number of infected humans. There was a decrease from 2,354 human cases reported in 2009 to 504 in 2010 (Dutch Ministry of EL&I, 2010; Van der Hoek et al., 2010; Roest et al., 2011; RIVM, 2012).

The outbreak of Q fever and the following control measures had a large impact on the society, especially on the goat milk industry. Because around 62,500 dairy goats were culled, there was a significant drop in goat milk production in the Netherlands. Farmers lost a large part of their income and the government was forced to spend a lot of money on this outbreak for compensation and control costs. In the goat milk industry, however, not only the farmers were at risk of being affected economically. This industry has a lot of other actors, including animal feed companies, transporters, dairy processors, meat goat industry, slaughterhouses and distributors. Many of these actors are directly linked to the farmers and had to adjust their production levels because of the decrease in number of animals in the industry. It was, however, not clear what the economic consequences for these actors had been.

A valuable tool to analyse the socio-economic impact of animal diseases or the related control measures is value chain analysis. The value chain consists of actors involved in the production of a certain commodity and the transport to the consumers, also known as the supply chain, combined with all the money flows that are generated by the production and consumption of the commodity (Rushton, 2010). A value chain analysis provides information on the structure, the trade flows and all the relations between the involved actors of a livestock sector.

The value chain analysis focuses on four main areas (Kaplinsky & Morris, 2001; Rushton, 2009; Rich & Perry, 2011): the mapping of the value chain, understanding the governance in the described industry, the possibility for actors to upgrade in the chain and the distribution of value among the different actors.

- *Mapping the value chain*. All the actors involved in the production, distribution, processing and end-sales as well as their interactions are described (Rushton, 2009). Next, these actors and their interactions are mapped in a chain, creating a flow scheme of a certain commodity in a livestock sector.
- *Governance in the value chain*. In this step the following questions should be addressed: which involved actor controls the coordination of economic activities, transactions and standards? Who has the most influence on the processing of the commodity and the distribution of profit between the different actors? (Humphrey & Schmitz, 2001; Rushton, 2009).
- *Upgrading in the value chain*. The possibilities for certain actors to improve their positions in a value chain can be studied. There are three different ways an actor can undertake changes for upgrading purposes: making the processing more efficient, making more sophisticated products and acquire new functions in the chain (Humphrey & Schmitz, 2000).
- *Distributional issues in the value chain*. The distribution of value among the different actors in the chain can be studied and quantified. In addition the benefit of the involved stakeholders to participate in this particular value chain should be assessed.

The value chain analysis describes the complex relations of the stakeholders and all the interactions that can be found in the livestock sector and how these connections change under the influence of an animal disease outbreak. The contemporary value chain analysis, however, often remains qualitative and descriptive (Rich & Perry, 2011; Rich et al., 2011). This makes a solid judgement of the economic impact for individual actors very difficult. In this study, the value chain of Dutch dairy goat products was quantified, resulting in a unique opportunity to quantify the economic impact of measures to control Q fever.

MATERIALS AND METHODS

The four main areas of the value chain analysis, as described above, were used as a guideline in this study.

Data collection

All the different groups of actors of the Dutch dairy goat industry were identified and a preliminary map of the value chain of dairy goat products was made. Although this industry relies heavily on the export of goat milk and other (intermediate) dairy products such as cheese and curd, only the part of the industry in the Netherlands was described. All of the activities outside of the Netherlands were categorised as export and not further described.

Value chain analyses can be executed in different ways with regard to the way actors are consulted for information about their enterprises and the concerning livestock industry. This can be done with group workshops and individual interviews (FAO, 2011; Rich et al., 2011).

The latter was chosen in this study to provide conditions under which people would be willing to share sensitive information.

The actors of the Dutch dairy goat industry were categorised in 20 different groups, and per group a different number of actors were individually interviewed, resulting in a total of 34 semi-structured interviews which were held in 2011. During these interviews questions were asked about their total revenues and gross margins, governance and upgrading possibilities in the value chain and the Q fever outbreak. These questions were asked in relation to 2009 and 2010. The impact of the majority of the control measures, taken from the end of 2009 until spring 2010, was measured using the economic situation in 2009 as a benchmark. Even though not all the groups were represented during the interviews, a major part of the most influential actors did participate and provided enough information for a valid analysis.

Complementary secondary data were obtained from interviews with agricultural accountants, the Dutch Federation of Agriculture and Horticulture and the Dutch Dairy Board, and from institutes including Statistics Netherlands and the Dutch Agricultural Economics Institute. They provided additional information about the average balance sheets of Dutch goat farmers, reports of damage of Q fever-affected farmers and economic characteristics of the Dutch dairy goat industry. Information on consumer prices of dairy goat products was obtained by visiting supermarkets and Islamic butcheries. These prices were corrected for the change in food prices over 2009, 2010 and 2011, provided by Statistics Netherlands.

Next, all information collected from the interviews and the secondary data was used to construct the final map of the value chain, as shown in Fig. 1. In the middle of the map (in grey boxes), the most important parts of the value chain of Dutch dairy goat products are shown. From the goat farmers, the goat milk passes by several people or organisations still as bulk milk or as dairy products, until it is exported or consumed in the Netherlands. Primary dairy processors are processors who collect milk directly from farmers, whereas the secondary dairy processors collect milk from milk collectors or other processors. Goat farmers depend on service providers, for example the veterinarians and artificial insemination stations, and the suppliers of feed, milk powder and milking systems. They also trade manure and kids (mostly bucks). The latter are sold for the production of meat. Even though these parts can be described as separate value chains, the part of the trade in kids and meat processing is too closely involved to the dairy part of the chain to be left out of the study. The goat value chain is supported by a number of people and organisations that facilitate movement and transactions. These are shown in the left part of the diagram. The value chain was complemented with the amount and prices of products and the money flow.

Data analysis

Value chain analysis: With the use of the value chain all the relations in the Dutch dairy goat industry were illustrated. The information of these relations was used to complement the value adding process of the chain. When actors were not willing to cooperate in this study, for example retailers, their selling and purchase prices could be established via the actors they were in business with. The value chain of 2009 was compared to the value chain of 2010, after the control measures of the Q fever outbreak were executed.

Gross margin analysis: The gross margin is the difference between an enterprise output and the variable costs. The gross margins of different actors can only be compared when the production systems are alike (Rushton, 2009). In this study, a lot of different companies participated with different production systems, therefore only the differences in gross margins between 2009 and 2010 were calculated per group of actors, not between the groups. To establish the gross margins of the retail, information was used from known direct relations between large retail entrepreneurs and dairy processors, and between Islamic butcheries and meat processors. With the provided purchase price and the consumer price (excluding tax) the gross margins of the different dairy and meat products of the retail could be calculated.

No information was available about the fixed costs of the actors. Therefore, it was not possible to calculate the net profit margins, or profit, of the participating actors. The output of the analyses was the percentage-wise difference between gross margins in 2009 and 2010 for the different groups and across the entire sector. Hence, the economic impact per group of actors could be established, when often only the impact of a whole industrial sector is shown.

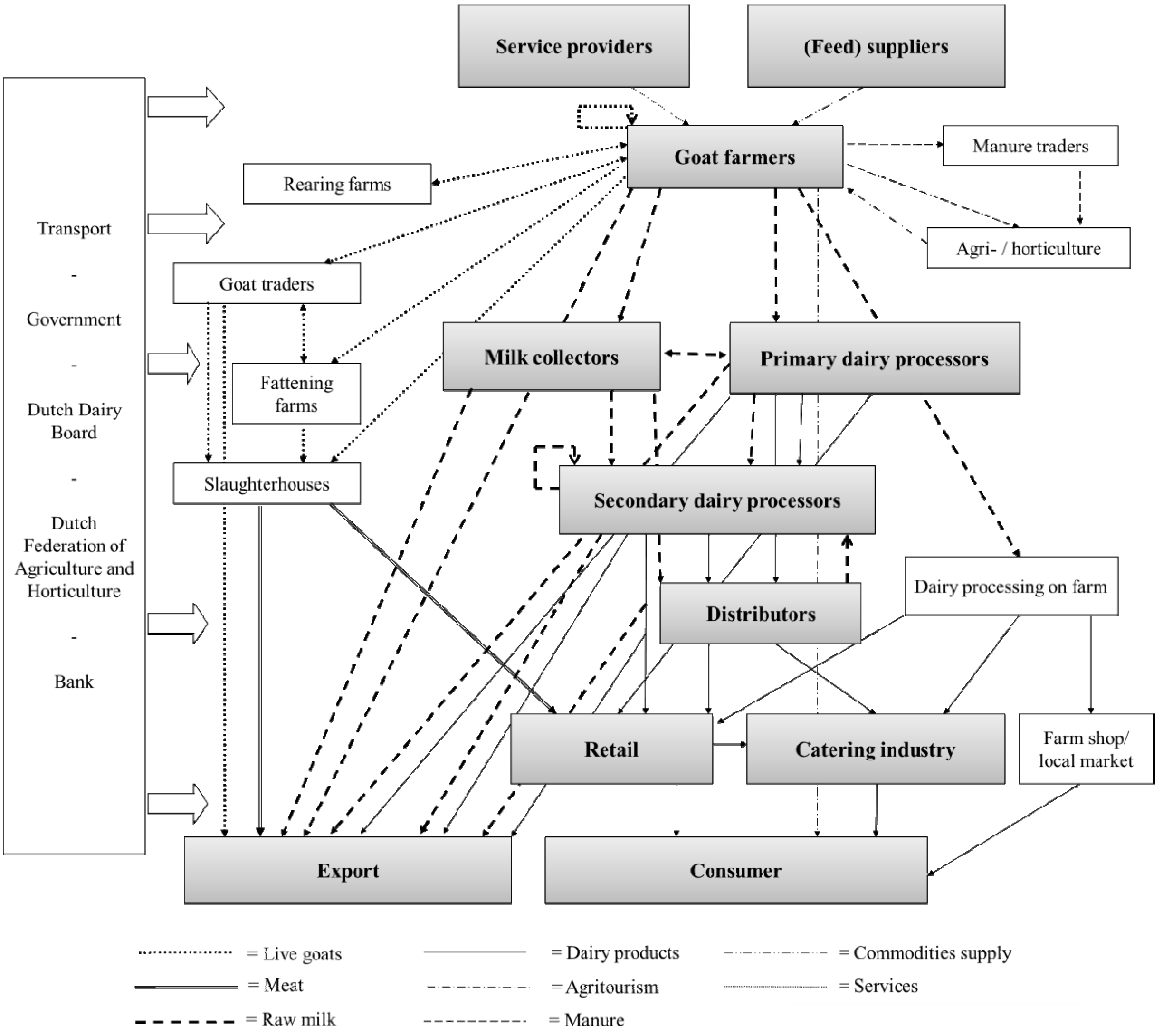


Fig. 1 The value chain of Dutch dairy goat products

RESULTS

Interviews

During all of the semi-structured interviews questions were asked about, among others, the governance in the chain. Various opinions were expressed but the retail and dairy processors were mentioned most frequently as the most powerful actors, followed by the distributors and milk collectors. This governance was often mentioned in relation to milk price-setting. Among the dairy processors, retail was most often mentioned as the group with most governance, followed by dairy processors themselves. They felt that they control and standardise the production process in the chain. Goat farmers were thought to be the group of actors who have the most opportunity to improve their position and to gain a better milk price. For this to happen, other actors suggested that farmers and their cooperations should collaborate more and stand together against the dairy processors and traders.

Value chain analysis

All the relations in the value chain remained intact between 2009 and 2010, except for one relation between a milk collector and a dairy processor, because of the inability of the milk collector to provide the dairy processor with enough milk. The demand of Dutch consumers for dairy goat products was not influenced by the Q fever outbreak. They did not associate the products with the disease even though the media brought a lot of attention to the Q fever outbreak.

Apart from one primary dairy processor, the actors in the main part of the value chain benefited from participation in the chain. In the peripheral parts, however, this could by no means be taken for granted. In previous years the main actors related to the buck trade and meat processing part of the chain did not profit from participation.

With all the information provided by the different actors, the amounts of milk and milk equivalents distributed through the chain in 2009 were made visible in Fig. 2. This figure distinctively shows the huge difference in distribution of milk to export and Dutch consumers.

Gross margin analyses

The gross margins of the farmers were calculated with information provided by two different accountants. This information was based on 72 farmers associated to these accountants. Damage reports of the Dutch Federation of Agriculture and Horticulture were used to estimate the gross margins of Q fever-affected farms in 2010. In 2009, the gross margin of the total number of goat farmers represented the largest part (45%) of the total gross margin of the industry, followed by the retail (28%). Between individual actors in the group of primary dairy processors as well in the group of secondary dairy processors, large differences in gross margins were found in 2009. The largest difference in gross margins between two primary dairy processors was 0.53 euro/kg, whereas between two secondary dairy processors the largest difference was 0.77 euro/kg.

In Fig. 3 the differences in gross margins between 2009 and 2010 are shown. The group of meat processing actors includes the goat traders, fattening farms, meat processor and slaughterhouses. The service providers were not included in the gross margin analyses since most of them were not able to provide these margins because of the large variety of services in their enterprises.

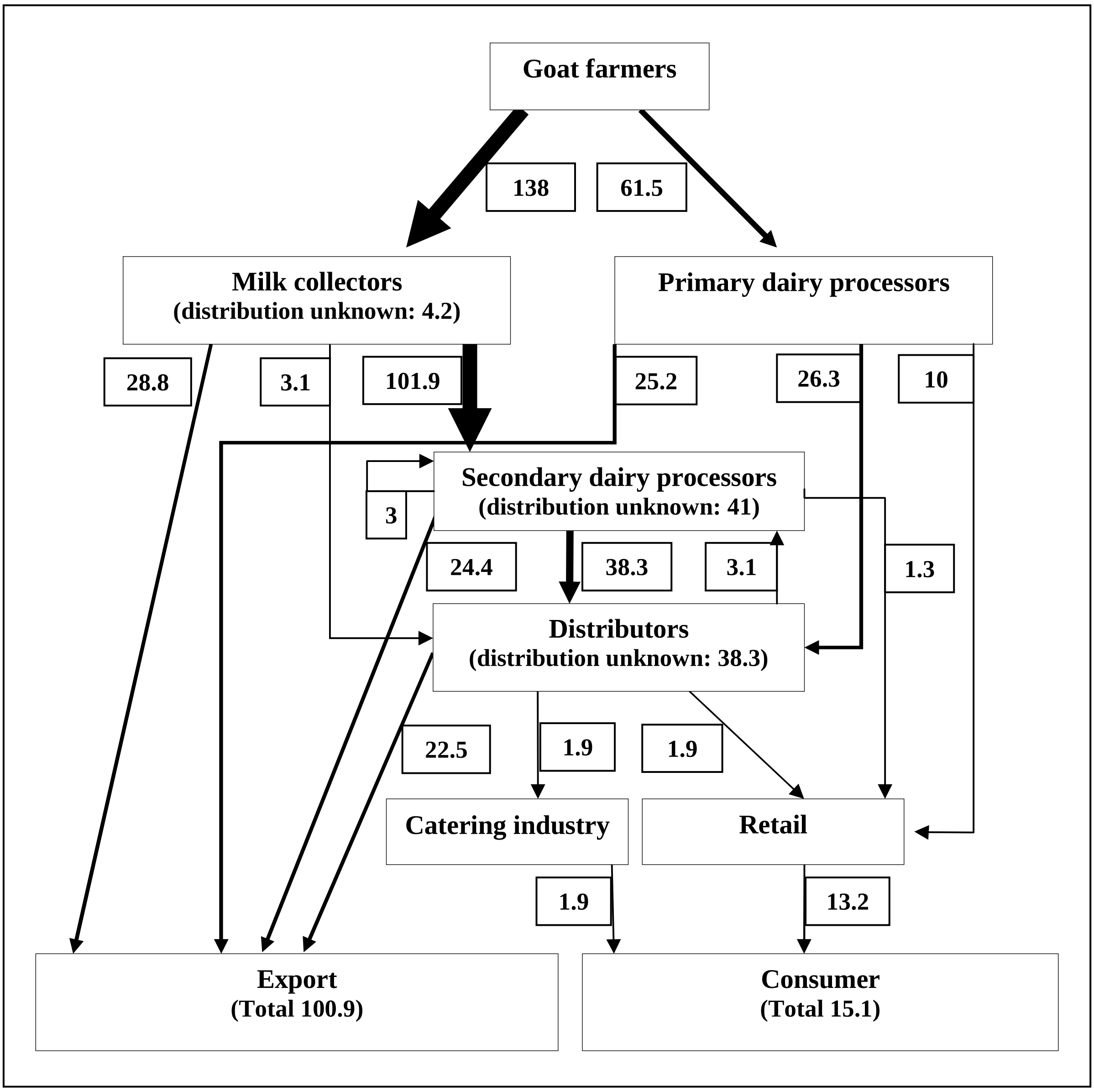


Fig. 2 The distribution of milk and milk equivalents in the Dutch dairy goat sector in 2009, in million kg milk. (Milk equivalents: hard cheese 1:10, soft cheese 1:6, yoghurt 1:1.2)

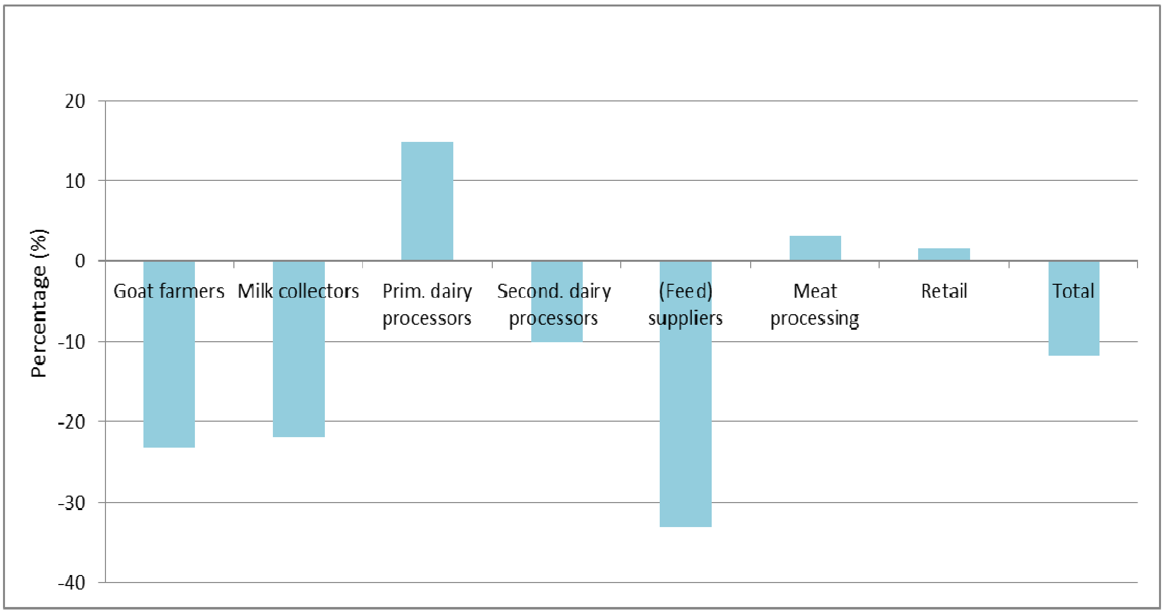


Fig. 3 The difference in gross margins in % between 2009 and 2010

The impact of Q fever control measures: At the end of December 2009 the most influential control measures, including the culling of pregnant goats, were executed to bring the increase of human Q fever cases to a halt. To study the economic effects of these measures, the distribution of value in the value chain of Dutch dairy goat products in 2010 was measured and compared to the value chain of 2009. This comparison was carefully chosen since the Dutch goat industry was still in development and experienced an average increase of milk production of 10% per year since 2000 (Dutch Dairy Board, 2012). Comparing the financial situation of 2010 to years before 2009 would have led to too many incorrect assumptions.

The overall decrease of gross margins of farmers in 2010, compared to 2009, was 10 million euros (-23%). There was, however, an enormous difference in decrease between the gross margins of Q fever-affected farms (-53%) and non-affected farms (-12%).

The milk collectors, secondary dairy processors and (feed) suppliers also suffered from the decrease in number of goats in the chain; their gross margins decreased with respectively 22%, 10% and 33% compared to 2009. The meat processing part of the chain, the retail and the primary dairy processors were not negatively affected by the Q fever control measures; compared to 2009 their gross margins increased with respectively 3%, 2% and 15%. The total gross margin of the value chain decreased with 11 million euros (-12%) (Fig. 3).

DISCUSSION

Quantification of the value chain

The main focus of the study was to quantify the economic impact of an animal disease and its related control measures on a livestock industry, using a value chain analysis. As a subject, the Q fever outbreak in the Netherlands was chosen because this outbreak was clearly constricted to one well defined livestock sector. The information provided by different actors of the value chain during individual interviews was the main source of data for this study. Because of the individual setting, people were willing to share sensitive information like the total revenues and gross margins of their enterprises. It was, however, not certain if the provided information was correct. During the analysis process, it appeared that not all of the actors provided accurate numbers, but rather rough estimations. The results of this study can therefore not be used to show the actual amount of euros lost due to the impact of Q fever on the dairy goat industry, only an estimation of this amount.

Unfortunately not all of the actors from the main part of the value chain were interviewed. This was often because of lack of motivation to participate. Especially the actors closest to the consumers, like dairy traders and the larger retail entrepreneurs were hard to motivate. This is a known problem of these actors in this type of studies (Te Velde et al., 2006) and can be explained by the assumption that their businesses were not affected by the outbreak of Q fever. Another explanation is the negative reputation concerning the profit these enterprises make per product.

The study shows the possibility to quantify the impact of an animal disease with the use of a value chain analysis. This could only be accomplished with the cooperation of the actors. Often value chain analyses are executed with the use of workshops with several actors of an industry. Although these workshops are very useful for discussions, actors will not reveal sensitive information about their enterprises in these settings, nor will they share controversial opinions about the industry and (relations of) other actors. Depending on the desired outcome of a value chain analysis, one should choose the right setting to gather information from actors. With the use of a value chain all the relations in the dairy goat industry were visualised. This was very useful to show where information on gross margins was missing and if this information could be gathered from other known relations with associated values.

Assumptions

The outbreak of Q fever among Dutch goats started in 2005, but the main control measures were not taken until the end of December 2009. The effects of these control measures on the dairy goat industry were clearly visible in 2010. The economic situation in this year was therefore chosen to measure the direct financial impact of the outbreak and the associated control measures. Because this livestock industry was still growing and developing in December 2009, the most recent economic situation of the industry, the year 2009, was used as a comparison for 2010. It is however not clear if the changes in the economic situation in 2010 were solely due to the effect of the outbreak and control measures.

As compensation for the culling of their goats, farmers received an amount of 250-600 euros per goat. The gross margins of the farmers were not directly affected by this compensation, because it was assumed this money was used to pay off credit loans and compensate for the devaluation of the reduced herd of goats or to buy new dairy goats.

Even though the retail did not participate in this study, their gross margins could be calculated using information from dairy processors directly related to the retail. The gross margins of the distributors could not be calculated because of lack of information of this part of the chain.

The impact of Q fever

As a consequence of the control measures taken to bring the outbreak of Q fever to a halt, the number of dairy goats in the Netherlands was significantly lowered. The affected farmers lost a large part of their revenues because of the declined production of milk. Their variable costs, however, decreased, for example feed costs, but this could not compensate for the loss of revenues. The gross margins of the milk collectors, secondary dairy processors and (feed) suppliers were also negatively affected by the reduction of number of animals in the value chain. Control programmes that significantly reduce the number of animals in an industry should therefore only be implemented when strictly necessary and with great caution. Even though farmers are most visibly affected by these kinds of measurements, other actors may face an even larger negative economic impact. Compensating only farmers for the economic damage caused by the implementation of control programs is therefore questionable.

ACKNOWLEDGEMENTS

This study could not have proceeded without the participation of actors of the Dutch dairy goat industry. All the participating actors of the following groups are therefore gratefully acknowledged for their inputs: goat farmers, milk collectors, dairy processors, feed companies, milking system suppliers, milk powder suppliers, veterinarians, A.I. station, semen inseminators, claw carers, accountants, animal health service, dairy traders, goat traders, meat processors, slaughterhouses, consumers, Dutch Dairy Doard and Dutch Federation of Agriculture and Horticulture.

REFERENCES

Bunte, F.H.J., Kuiper, W.E., van Galen, M.A. and Goddijn, S.T. (2003). Macht en Prijsvorming in Agrofoodketens. Report 5.03.01, LEI, Den Haag, The Netherlands

Dutch Agricultural Economics Institute (LEI) (2012). Data on Dutch dairy goat sector. http://www3.lei.wur.nl/ltc/Classificatie.aspx. (accessed January 24th, 2012)

Dutch Dairy Board (Productschap Zuivel) (2012). Market, production and price data. http://www.prodzuivel.nl/ (accessed on 24.01.2012)

Dutch Federation of Agriculture and Horticulture (2010). Schaderapport Q koorts. February 4, 2010, Tilburg, The Netherlands

Dutch Ministry of Economic Affairs, Agriculture and Innovation (EL&I) (2010). Van verwerping tot verheffing: Q koorts beleid in Nederland 2005-2010. November 22, 2010, Den Haag, The Netherlands, found at http://www.rijksoverheid.nl/documenten-en-publicaties/rapporten/ 2010/11/22/van-verwerping-tot-verheffing.html

FAO (2011). A value chain approach to animal diseases risk management – Technical foundations and practical framework for field application. Animal Production and Health Guidelines. No. 4. Rome, Italy

Humphrey, J. and Schmitz, H. (2000). Governance and upgrading: linking industrial cluster and global value chain research. IDS Working paper 120, Institute of Development Studies, 2000

Humphrey, J., and Schmitz, H. (2001). Governance in global value chains. In: IDS Bulletin 32.3, Institute of Development Studies, 2001

Kaplinsky, R. and Morris, M. (2001). A Handbook for Value Chain Research. Working Paper Prepared for the IDRC, Institute of Development Studies, Brighton, UK

McLeod, A., Kobayashi, M., Gilman, J., Siagian, A. and Young, M. (2009). The use of poultry value chain mapping in developing HPAI control programmes. Worlds Poult. Sci. J. 65, 217-224

Rich, K.M. (2007). New methods for integrated models of animal disease control. The American Agricultural Economics Association Annual Meetings, Portland, OR, 31 July 2007, http://ageconsearch.umn.edu/handle/9701 (accessed January 22nd, 2013)

Rich, K.M. and Perry, B.D. (2011). The economic and poverty impacts of animal diseases in developing countries: New roles, new demands for economics and epidemiology. Prev. Vet. Med. 101, 133-147

Rich, K.M., Ross, R.B., Baker, A.D. and Negassa, A. (2011). Quantifying value chain analysis in the context of livestock systems in developing countries. Food Policy 26, 214-222

RIVM (National Institute for Public Health and the Environment) (2012). http://rivm.nl /Onderwerpen/Ziekten_Aandoeningen/Q/Q_koorts. (accessed January 24th, 2012)

Roest, H.I.J., Tilburg, J.J.H.C., van der Hoek, W., Vellema, P., van Zijderveld, F.G., Klaassen, C.H.W. and Raoult, D. (2011). The Q fever epidemic in the Netherlands: history, onset, response and reflection. Epid. Inf. 139, 1–12

Rushton, J. (2009). The Economics of Animal Health and Production. CABI Publishing, Wallingford, UK

Rushton, J., Taylor, N., de Haan, N., Huntington, B., White, L., Kalpravidh, W. and Rushton, R. (2010). Value Chain Analysis. In: Proceedings of the Dutch Society for Veterinary Epidemiology and Economics, Lelystad, 23-25 November

Statistics Netherlands (CBS). (2012). Data on Dutch dairy goat sector. http://statline.cbs.nl/StatWeb/publication/?DM=SLNL&PA=80780NED&D1=440-450,482,486-487,490&D2=0&D3=0,5,(l-2),(l-1),l&HDR=G1,G2&STB=T&VW=T. (accessed January 24th, 2012)

Te Velde, D.W., Rushton, J., Sckreckenberg, K., Marshall, E., Edouard, F., Newton, A. and Arancibia, E. (2006). Entrepreneurship in value chains of non-timber forest products. Forest Pol. Econ. 8, 725–741

Van der Hoek, W., Dijkstra, F., Schimmer, B., Schneeberger, P.M., Vellema, P., Wijkmans, C., ter Schegget, R., Hackert, V. and van Duynhoven, Y. (2010). Q fever in the Netherlands: an update on the epidemiology and control measures. Eurosurveillance 15, 25 March 2010

OBSERVATIONAL STUDIES

A CROSS-SECTIONAL STUDY OF *SARCOCYSTIS* SPP. IN LLAMAS (*LAMA GLAMA*) IN SOUTHERN BOLIVIA INCLUDING AN ECONOMIC COSTING OF LOSS IN INCOME CAUSED BY THE DISEASE

A.L. ROONEY[*], G. LIMON, H. VIDES, A. CORTEZ AND J. GUITIAN

SUMMARY

Llamas (*Lama glama*) are intermediate hosts of the protozoan parasite *Sarcocystis* spp. This parasite is described as causing economic losses in the production of llama meat in South America. The aim of this study was to estimate prevalence, identify risk factors and explore any spatial patterns of *Sarcocystis* in llamas in an area of the Bolivian High Andes including estimating some economic losses due to the disease in this area.

Overall this study revealed a high prevalence of *Sarcocystis* in the study area with some heterogeneity between districts. It also identified some previously unknown risk factors for *Sarcocystis* and gave some economic estimates of the cost of the disease. It is hoped these findings will add to the understanding of *Sarcocystis* in llamas in Southern Bolivia and will be useful when considering if controls are necessary, worthwhile and practical for this parasite.

INTRODUCTION

Llamas are an important source of income for the people of the Bolivian High Andes by provision of protein, hide, fibre and transport (Leguia, 1991). Llama meat is often regarded as a low value product compared to other protein sources such as beef or mutton, mainly due to the presence of sarcocystiosis (Viscarra et al., 2002). Sarcocystiosis is caused by the protozoan parasite *Sarcocystis* of which two species are identified in llamas, *S.aucheniae* and *S.lamacenis* (Taylor et al., 2007). *Sarcocystis* is a member of the *Sarcocystidae* family of protozoan parasites which also include the genera *Toxoplasma* and *Neospora*. The life cycle of *Sarcocystis* spp. is similar to that of *Toxoplasma* and *Neospora* with the involvement of an intermediate and final host. The intermediate host is typically a herbivore or omnivore such as a cow, sheep or pig and the final host is usually a predator species such as the dog or cat. The parasite passes through the intestines of the final host where sporocysts, infective to the intermediate host, are passed in faeces. The intermediate host then picks up these sporocysts

---

[*] Ann Rooney, Basement Flat, 14 St Lawrence Terrace, London, W10 5SX, UK.
Email: annlouiserooney@gmail.com

by ingestion. The parasite encysts in muscle cells in the intermediate host and this is the infective stage for the carnivorous final host, passed on by ingestion (Urquhart et al., 1996). It is thought that most *Sarcocystis* species are species-specific for intermediate hosts and family-specific for final hosts (Tenter, 1995). Of the *Sarcocystis* species found in llamas it is thought that *S.aucheniae* uses the dog as a final host, however the final host for *S.lamacenis* is not known (Taylor et al., 2007). There has been little research undertaken to investigate if any of the *Sarcocystis* species found in llamas are infective to humans. Some species such as *S.hominis* and *S.suihominis* are known to include man as a final host (Fayer, 2004). There is some suggestion of zoonotic potential of *Sarcocystis* spp. in llamas (Leguia, 1991) but evidence is limited and it is not widely known if the *Sarcocystis* spp. harboured in llamas is transmissible to humans (Fayer, 2004). Humans would typically contract infection with *Sarcocystis* by ingestion of raw or undercooked meat products containing mature sarcocysts. Infection of llamas with *Sarcocystis* can cause important economic loss due to downgrading of meat infected with the parasite (Leguia, 1991; Tenter, 1995; More et al., 2008; Fowler, 2010). Meat with visible cysts attributable to *Sarcocystis* is deemed unfit for human consumption unless it is processed into a dehydrated form of meat termed ‘charque’. This is an effective way of destroying sarcocysts (Leguia, 1991) ensuring meat can be sold for human consumption; this meat can however have a lower value compared to fresh meat.

A report investigating the validation of a serological test in llamas for *Sarcocystis* in Bolivia estimated prevalence on meat inspection to be 35% (Viscarra et al., 2002). Other studies have estimated a seroprevalence of 96% in llamas in Argentina (More et al., 2008) and 89.7% (95% confidence interval 87.8, 91.6) in alpacas (*Lama pacos*) in Peru (Castro et al., 2004). The alpaca study identified age as a risk factor where it was shown that animals over 1 year old had a significantly higher seropositivity for *Sarcocystis* compared to animals below 1 year of age (Castro et al., 2004). Documented economic assessments of the impact of *Sarcocystis* in llamas are scarce. It has been estimated that sarcocystiosis makes up 11.4% of the total annual losses directly attributable to parasitic disease in alpacas in Peru (Leguia, 1991) and some work has been done in Bolivia (Meijer et al., 2000) from data collected from a number of different abattoirs throughout the country to report on the economic impact of *Sarcocystis* in llamas which showed economic loss in terms of downgrading of meat.

This study used data that were gathered in a newly established llama abattoir in the Bolivian High Andes between 2006 and 2011 to answer specific research questions about the epidemiology and impact of *Sarcocystis* in this area, outlined by the following study objectives: to estimate the prevalence of *Sarcocystis* in llama carcasses; to identify risk factors for *Sarcocystis* infection; to describe and explore the spatial distribution of *Sarcocystis* in the area; and to quantify the economic loss due to *Sarcocystis* by the parasite causing downgrading of meat.

MATERIALS AND METHODS

Data used in this study were collected on every llama slaughtered in an abattoir recently established specifically for the slaughter of llamas in Southern Bolivia. Origin of the animal, age, sex, type (analogous to breed), colour, live weight, carcass weight, slaughter date, price per kilogram live weight and destination of the meat were routinely gathered by a veterinary surgeon at the abattoir who performed meat hygiene and animal welfare duties. In total, observations for 1196 llamas had been recorded. Any meat found to have 1 or more sarcocysts at meat inspection would be required to be sold as dehydrated meat therefore all meat recorded to be destined ‘charque’ (dehydrated meat) was defined as from *Sarcocystis* positive animals at the point of slaughter. All meat termed ‘venta’ (fresh meat) was defined as from *Sarcocystis* negative animals at the point of slaughter, as these animals had no evidence of cysts.

Prevalence estimates of *Sarcocystis* were obtained for each individual year and for all 6 years combined to obtain a period prevalence between 2006 and 2011. 95% confidence intervals (CI) were obtained. Univariate associations between potential risk factors for *Sarcocystis* status were explored and adjusted associations were assessed by means of mixed effects logistic regression allowing for the correlation of observations by owner and community as random effects. Observations for all 6 years were combined to maximise power. Models excluding risk factors with little evidence of being associated with *Sarcocystis* ($P<0.05$, Wald test) were checked, both with and without interaction, to assess collinearity between risk factors. The reliability of the parameter estimates from both models were checked by varying the number of integration points of the model. All statistical analysis was carried out in Stata 12, StataCorp LP.

The spatial pattern of disease was explored by calculation and display of standardised morbidity ratios (SMRs), associated standard errors and 95% CI. SMRs where calculated using the observed number of *Sarcocystis* positive animals per district for all 6 years combined divided by the expected number of *Sarcocystis* positive animals per district for all 6 years combined.

To quantify economic losses due to *Sarcocystis* the total income for all owners per year if all the meat had been *Sarcocystis* negative was estimated by calculating a price per kilogram live weight that would have been given if all the *Sarcocystis* positive animals were negative.

RESULTS

Prevalence estimation

Prevalence per year varied from 23.4% to 50.3% between 2006 and 2011. The period prevalence of Sarcocystis between 2006 and 2011 was estimated at 34.1% (95% CI 31.4, 36.8).

Risk factor analysis

The results of the mixed effects logistic regression model allowing for clustering by owner and community showed that age, sex and type were risk factors for *Sarcocystis*. The odds of a llama in the oldest age group of 4.5 years or more being *Sarcocystis* positive was 19 times that compared to the youngest age group of 1-2.49 years (OR 19.31, 95% CI 9.10, 40.98, $p<0.001$). The odds of a female llama being *Sarcocystis* positive was almost twice that compared to a male llama (OR 1.75, 95% CI 1.13, 2.68, $p=0.01$) and the odds of a llama of long hair type being *Sarcocystis* positive was almost twice that compared to a intermediate hair type llama (OR 1.80, 95% CI 1.26, 2.87, $p=0.002$).

An interaction was detected during the analysis between age and sex in the age category 2.5-4.49 years: the odds of a female in the age category 2.5-4.49 years being *Sarcocystis* positive was almost 3 times that compared to the youngest males (OR 2.77, 95% CI 1.73, 4.42, p<0.001) whereas male llamas showed no evidence of being any more likely to be *Sarcocystis* positive when compared to young males (OR 1.16, 95% CI 0.80, 1.70, p=0.433).

Spatial patterns

Table 1 shows the distribution of Sarcocystis positive and negative llamas by district. This initial description shows some heterogeneity between districts with regard to the number of Sarcocystis positive llamas.

Table 1. Distribution of *Sarcocystis* positive (n=397) and *Sarcocystis* negative (n=767) llamas between 2006 and 2011 by district

District	Number of carcasses *Sarcocystis* positive (%)	Number of carcasses *Sarcocystis* negative (%)	Total (%)
Copacabana	314 (36.8)	539 (63.2)	853 (100.0)
Quebrade Honda	49 (23.6)	159 (76.4)	208 (100.0)
San Luis del Palqui	5 (100.0)	0 (0.0)	5 (100.0)
Yunchara	3 (60.0)	2 (40.0)	5 (100.0)
Tojo	6 (9.8)	55 (90.2)	61 (100.0)
Curqui	18 (78.3)	5 (21.7)	23 (100.0)
Iscayachi	2 (22.2)	7 (77.8)	9 (100.0)

Figures 1a and 1b show area maps of SMRs at district level (Fig. 1a) and associated standard errors (Fig. 1b) for Sarcocystis. This shows a geographic difference in the location of the districts showing heterogeneity for risk of Sarcocystis. Curqui is located in the north of the area whilst Quebrada Honda and Tojo are located in the west and south east respectively.

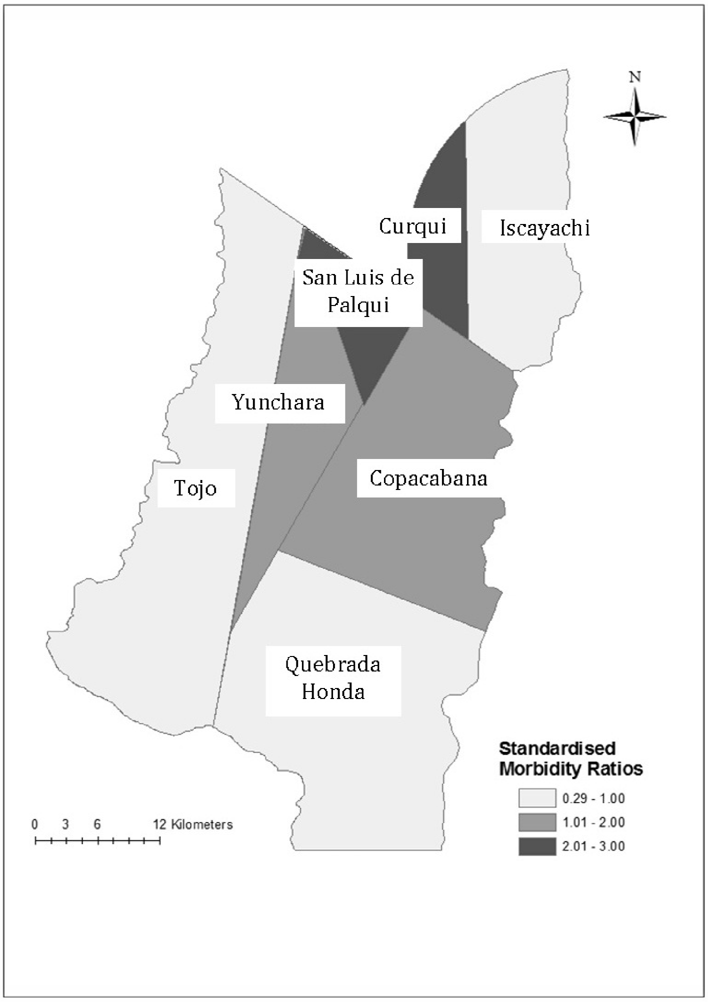


Fig. 1a District map showing standardised morbidity ratios (SMRs) for *Sarcocystis* in Municipalities of Yunchará and part of El Puente in Bolivia

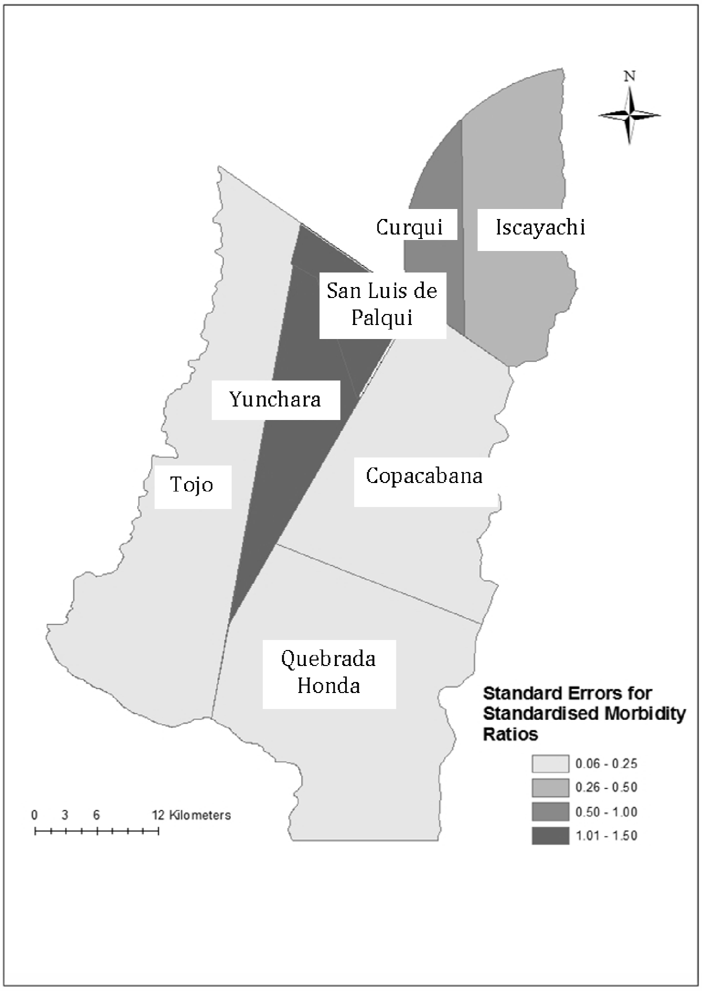


Fig. 1b District map showing standard errors for the standardised morbidity ratios for *Sarcocystis* in Municipalities of Yunchará and part of El Puente in Bolivia

Figure 2 shows the distribution of llama density by district. This illustrates that the districts of Curqui, Quebrada Honda and Tojo identified as showing a difference in risk of *Sarcocystis* do not show difference in llama density. All 3 districts have between 1.01 and 4.99 llamas per square kilometre.

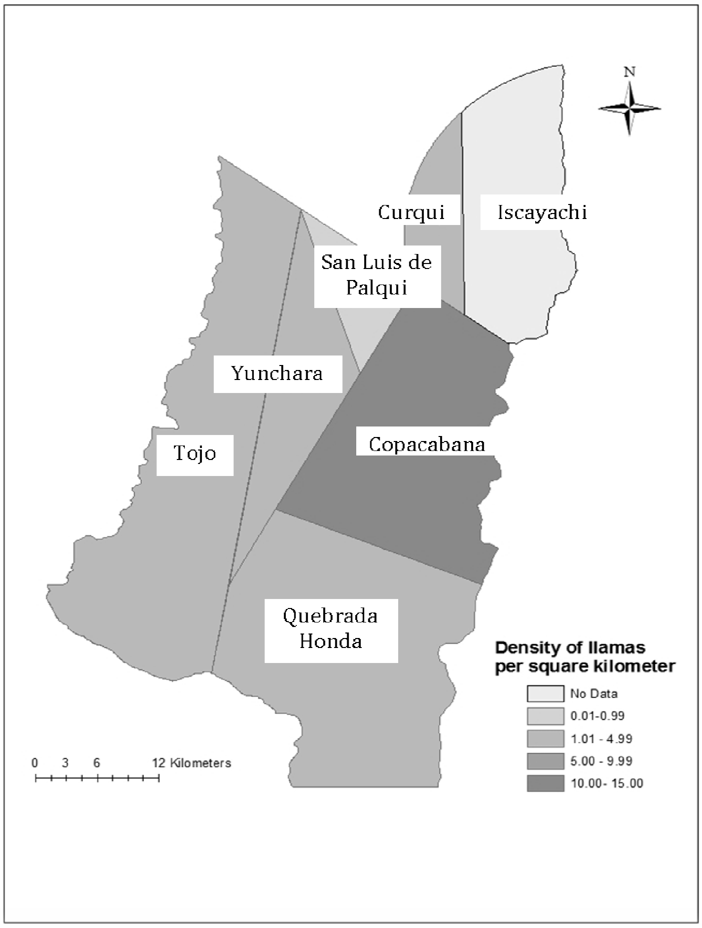


Fig. 2 District map showing distribution of llama density by district in Municipalities of Yuncharáand part of El Puente in Bolivia

Economic costing

In the earlier years of the study period (2006 and 2007) the loss remained under 10% of the actual income but in the later years (2008 to 2011) the loss as a percentage of actual income rose, varying between 13.2% in 2010 and 20.6% in 2009.

DISCUSSION

The high prevalence estimated in this study is in agreement with a prevalence estimate from previous work also carried out in the Bolivian High Andes (Viscarra et al., 2002). Reasons for high prevalence of *Sarcocystis* in the area are not known. Anecdotally dogs and llamas are said to roam free in all districts, which may aid in the spread of the parasite; in addition llamas are commonly slaughtered at home and the uncooked offal fed to dogs (Viscarra et al., 2002). It should also be noted that the final host for *S. lamacenis* is not known and without this the full reasons for the maintenance of the parasite life cycle can not be explored.

The difference in risk of *Sarcocystis* between districts identified in this study could be due to factors such as density of final hosts or differences in environmental conditions such as temperate, altitude or humidity between districts.

Age, which had previously been described as a risk factor for *Sarcocystis* infection in alpacas (Castro et al., 2004), is a risk factor also in llamas, with older animals more likely to have detectable disease than younger animals. The effect of age on the likelihood of a llama being *Sarcocystis* positive was not the same between sexes with only females being at increased risk of disease between 2.5-4.49 years of age. The increased risk of *Sarcocystis* in long hair type llamas could be due to genetic factors making this type more susceptible to infection.

Considerable monetary losses are incurred due to the presence of *Sarcocystis* in meat. This is in agreement with previous work carried out in Bolivia to assess the economic impact of this disease (Meijer et al., 2000). It should be taken into consideration that the estimated economic costings do not take account of issues such as market supply and demand for fresh or dehydrated forms of llama meat. These estimates do however give some indication towards how much could potentially be saved if the disease were not present in this area, and although complete eradication of disease is not a practical target, these estimates could help toward making an assessment as to whether control of this parasite is worthwhile.

When interpreting our findings, it is important to bear in mind that *Sarcocystis* positive status was defined on the basis of the detection of one or more cysts during meat inspection, deeming meat ‘charque’. It is unknown how thoroughly each carcass was checked and if only certain muscles groups were targeted for examination. Sarcocysts in muscle are also very small, so for both these reasons false negative carcasses are a possibility.

In conclusion, *Sarcocystis* infection is highly prevalent among llamas in this area of the Bolivia High Andes, and its distribution shows some spatial heterogeneity. Age, sex and type of llama are risk factors for the infection with a difference in how age affects the risk in female and male llamas. We estimated that up to 20% of the annual income gained by the sale of llama meat in this area is lost because of the disease.

ACKNOWLEDGEMENTS

Thanks goes to all those involved with the abattoir in Bolivia for provision of the data and also the BBSRC for provision of a study scholarship for the main author during the period of this work.

REFERENCES

Castro, E.C., Sam, R.T., López, T.U., González, A.Z. and Silva, M.I. (2004). Evaluación de la edad como factor de riesgo de seroposividad a *Sarcocystis* sp. en Alpacas (in Spanish). Rev. Inv. Vet. Perú 15(1), 83-86

Fayer, R. (2004). *Sarcocystis* spp. in human infections. Clin. Microbiol. Rev. 17(4), 894-902

Fowler, M.E. (2010). Medicine and Surgery of Camelids. Wiley-Blackwell, Ames, Iowa, 251

Leguia, G. (1991). The epidemiology and economic impact of llama parasites. Parasitology Today 7(2), 54-56

Meijer, M., Rushton, J., Viscarra, R., Kay, D., Sonco, M. and McGrane, J. (2000). El impacto de la Sarcocistiosis en los sistema de comercializacion de carne de Llama en Bolivia (in Spanish). UNIVEP, Santa Cruz, Bolivia, 1-48

More, G., Pardini L., Basso, W., Marín, R., Bacigalupe, D., Auad, G., Venturini, L. and Venturini, M.C. (2008). Seroprevalence of *Neospora caninum*, *Toxoplasma gondii* and *Sarcocystis* sp. in llamas (*Lama glama*) from Jujuy, Argentina. Vet. Parasitol. 155, 158-160

Taylor, M.A., Coop, R.L. and Wall, R.L. (2007). Veterinary Parasitology. Blackwell Publishing, Oxford, UK, 47

Tenter, A.M. (1995). Current research on *Sarcocystis* species of domestic animals. Int. J. Parasitol. 25, 1311-1330

Urquhart, G., Armour, M.J., Duncan, J.L., Dunn, A.M. and Jennings, F.W. (1996). Veterinary Parasitology. Blackwell Publishing, Oxford, UK, 234-241

Viscarra, D.R., Rushton, D.J. and González, A. (2002). Validación de la prueba serólogica para Sarcocistiosis en llamas del Altiplano Boliviano (in Spanish).CEVEP, La Paz, Bolivia, 1-22

CANINE CHRONIC KIDNEY DISEASE IN UK VETERINARY PRACTICES:

PREVALENCE, RISK FACTORS AND SURVIVAL

D.G. O'NEILL[*], J. ELLIOTT, D.B. CHURCH, P.D. MCGREEVY, P.C. THOMSON
AND D.C. BRODBELT

SUMMARY

Analyses covering 107,214 dogs attending 89 UK practices aimed to estimate the prevalence and identify risk factors associated with canine chronic kidney disease (CKD) diagnosis and survival. The apparent prevalence (AP) was estimated using a cross-sectional approach while the true prevalence (TP) was estimated using Bayesian analysis. A nested case-control study design was used to evaluate risk factors. Survival analysis was conducted using the Kaplan-Meier survival curve method and multivariable Cox proportional hazards regression modelling.

CKD AP was estimated at 0.21% and TP at 0.37%. CKD risk factors included increasing age, being insured and certain breeds. The median survival time was 226 days. IRIS stage and blood urea nitrogen concentration at diagnosis were significantly associated with hazard of CKD death.

This study of a clinically important condition indicates the increasing importance of electronic patient records (EPRs) to evaluate diseases in companion animals.

INTRODUCTION

Concern has long been expressed over possible adverse welfare effects of pedigree dog breeding practices on modern dog breeds (Hein, 1963; Hodgman, 1963; Broom, 1991; McGreevy & Nicholas, 1999). However, the issue was dramatically brought to the public consciousness in 2008 with the airing of a BBC documentary that graphically illustrated the suffering of dogs afflicted by alleged breed-associated disorders such as syringomyelia and epilepsy (BBC, 2008). Three major UK reports followed that examined the canine health impacts of established dog breeding practices and all identified that current dog breeding practices did pose substantial welfare problems for dogs but concluded that a deficiency of accurate prevalence data on the common disorders affecting pedigree dogs was a limiting

---

[*] Dan O'Neill, Veterinary Epidemiology, Economics and Public Health, The Royal Veterinary College, Hawkshead Lane, North Mymms, Hatfield, Herts AL9 7TA, UK.
Email: doneill@rvc.ac.uk

factor for effective reforms (APGAW, 2009; Rooney, 2009; Bateson, 2010). These reports recommended the systematic collection of computerised anonymised diagnoses from veterinary surgeries as the optimal data source to provide robust prevalence data (Rooney, 2009) for common disorders among dog breeds.

In the UK, approximately 86% of primary veterinary practices use electronic practice management systems (PMSs) to record demographic, medical and treatment information (Robinson & Hooker, 2006) making veterinary electronic patient records (EPRs) a potentially vast and useful resource for companion animal health data. The majority of practices use commercial PMS packages offered by around 15 major PMS providers but there is increasing development of bespoke systems, especially among the corporate practice groups. There are estimated to be around 5,000 clinical practices in the UK employing around 16,000 veterinary surgeons (RCVS, 2012). A general public survey estimated that approximately 72% of dogs were vet-visiting in the UK (Asher et al., 2011). The UK dog population has been variably estimated at between 6 and 11 million with 75% of these being purebred (Thrusfield, 1989; Murray et al., 2010; Asher et al., 2011; PFMA, 2012) indicating the opportunity for substantial welfare benefits from increased understanding and effective reforms relating to inherited disease.

Chronic kidney disease (CKD) is the most prevalent kidney disorder of dogs (Polzin et al., 2005) and describes kidney disease that has been present for an extended period (3 months or longer), is irreversible and generally progressive (Polzin, 2011). However, estimated prevalence values have ranged widely, with previous work reporting values of 0.05% (Macdougall et al., 1986), 0.9% (Polzin et al., 2005), 0.5-1.5% (Brown, 2007) and 3.74% (Sosnar, 2003) depending on the study design used and the populations studied. Affected dogs are reported as surviving up to 2 years from diagnosis (Polzin, 2011). The CKD end-state is believed to result from a range of inherited, congenital and acquired factors (Nelson & Couto, 2009) and a number of breeds have been previously reported to be at increased risk of the disease, although not all these findings may be relevant to the UK canine population (Jones et al., 1989; DiBartola et al., 1990; Lees et al., 1998; McAloose et al., 1998). Hence, CKD has potential major welfare implications for dogs in general and specific breeds in particular and further data to evaluate its current frequency in the UK and factors associated with it would aid its diagnosis and management. This paper aims to describe the methodology and main findings of a study of CKD as an example of a disease with breed implications and major welfare consequences that could be studied using EPR data.

MATERIALS AND METHODS

Data collection

The VetCompass project shares companion animal EPR data from primary veterinary practices in the UK (VetCompass, 2012). The project was piloted from 2007-2009 among 7 clinics to refine and validate EPR extraction and upload before full project implementation commenced from 2010 onwards. The project will run long-term with progressively extending lifetime coverage for animals under observation as well as increasing practice and patient numbers over time.

The current CKD study included all EPR data from practices actively contributing to the VetCompass database during a 2-year period between 1st January 2010 and 31st December

2011. Practices were selected as a convenience sample on the basis of willingness to participate and their use of a compatible PMS and were situated mainly in central and south-eastern England. A lexicon of veterinary-specific terms (species, breed, procedure, summary diagnosis term) called the VeNom Codes was embedded into the PMSs of participating practices (The VeNom Coding Group, 2012). Attending veterinary surgeons were encouraged to select and record the most appropriate VeNom Code summary diagnosis term(s) at each episode of clinical care in addition to their routine recording of free-form text notes. An integrated clinical query (Upjohn et al., 2008) extracted specified de-identified EPR fields (Table 1) for automatic weekly upload to a password-protected file transfer protocol site.

Table 1. VetCompass data fields extracted from practice management systems

Identifier fields	Demographic fields	Clinical fields
Animal ID[a]	Species	Date of care
Owner ID[a]	Breed	Passport Yes/No
Clinic ID[a]	Sex	Insured Yes/No
Vet ID[a]	Desexed	Temperature
Visit ID[a]	Colour	Clinical Exam text
Microchip number	Weight	VeNom summary diagnosis
Partial postcode	Date of birth	Treatment
	Registration date	Deceased date

[a] ID Identification Number

A location-based secure structured query language (SQL) database was developed with an Access (Microsoft Access, Microsoft Corp) front-end to store the EPR data. Permanent variables were stored within a single ‘Patient’ table with additional individual tables storing non-permanent variable data (neuter, insurance, microchip, partial postcode, weight, clinical notes, summary term diagnoses, treatment). The Data Protection Act (1998) and Royal College of Veterinary Surgeons (RCVS) guide to professional conduct requirement were complied with. Ethics approval was granted by the Royal Veterinary College Ethics & Welfare Committee (URN 2011 1104).

Study design

Prevalence: A cross-sectional approach was used to estimate the prevalence of canine CKD where all dogs with records shared during the study period were included and evaluated for CKD diagnosis.

Risk factor analysis: A nested case-control study design was used to evaluate risk factors associated with CKD diagnosis. Random control dogs were selected from the overall VetCompass dataset sampling frame to attain a final 1:1 case-control ratio using a web-based random number generator (Haahr, 2012). Potential control animals with a history indicative of kidney disease were excluded. Sample size estimates were calculated in order to detect risk factors with an odds ratio of 2.0 or greater and assuming 15% prevalence in the control animals. The results indicated that approximately 209 cases and 209 controls would be required for 80% power with a 95% confidence level using an unmatched 1:1 case-control study.

Survival: CKD cases were followed forward in time within the clinical records to identify clinical outcomes (died of CKD, died of other cause, censored due to other reasons, alive at study termination).

Case definition and identification

Potential CKD cases were identified by querying the 'VeNom summary diagnosis' field dogs for appropriate VeNom diagnosis terms (chronic kidney (renal) disease, renal (kidney) disorder) and by searching the free-text clinical notes using SQL database queries that included the search terms: renal, kidney, CKD, CRD, CKF, CRF, azot*, urem* and uraem*. Case confirmation depended heavily on the primary practitioner's overall clinical acumen and resources for making the CKD diagnosis. Case inclusion criteria specifically required both i) a recorded summary diagnosis term, insurance claim term or veterinarian diagnosis of CKD with a consistent history and ii) evidence of blood biochemistry analysis to assist diagnosis. The diagnosis date related to the first confirmatory blood sample. Where diagnosis preceded available records, prior clinical histories (n=15) were sourced from participating practices. Clinical laboratory values (creatinine, urea, phosphate) at diagnosis and urinalysis results obtained closest to the diagnosis date (maximum 28-day window) were noted. Dates for the earliest and latest live animal EPR, all documented clinical signs and co-morbid disorders and, for animals dying during the study, the cause of death or stated reason for euthanasia were recorded.

Data management

Dog breed data were recorded in VeNom Code format offering 767 breed options. Breed data were further categorised to define 3 additional breed-related variables: purebred/crossbred, Kennel Club (KC) breed status (registered/not registered breed) and KC breed group status [gundog, hound, pastoral, terrier, toy, utility, and working] (The Kennel Club, 2012). Breeds with 12 or more dogs within the case-control study were evaluated individually. The weight, insurance and neuter status values at the final EPR were used. The age at the final live EPR was included for the case-control study while the age at diagnosis was used for survival analysis. Age (years) was categorised into 4 rounded quantiles; <4, 4-7, 7-12 and >12. Weight (kg) was categorised into 5 rounded quantiles: <7, 7-11, 11-20, 20-30, and >30. Blood creatinine concentration (mg/dl) at diagnosis was categorised following IRIS guidelines (IRIS, 2007); Stage 1: <1.4, Stage 2: 1.4-2.0, Stage 3: 2.1-5.0, Stage 4: >5.0 with Stage 1 and 2 collapsed for analysis (IRIS, 2007). Blood urea nitrogen (BUN) concentration values (mg/dl) at diagnosis were categorised into 4 groups: <44.8, 44.8 to <64.4, 64.4 to <112.0 and ≥112.0.

Data checking and cleaning was done in Microsoft Office Excel 2007 (Microsoft Corp.) and the majority of analyses were conducted in Stata Version 11.2 (Stata Corporation). The true prevalence (TP) estimation additionally used OpenBUGS version 3.2.1 rev 781.

Bayesian true prevalence analysis

Bayesian analyses were implemented in OpenBUGS version 3.2.1 rev 781 (Thomas et al., 2006; Thomas, 2009) and aimed to estimate TP based on a non-informative prevalence prior, the apparent prevalence (AP) estimated using Stata 11.2 and primary practice caseload diagnostic sensitivity (Se) and diagnostic specificity (Sp) values. No published Se and Sp information for CKD was available so expert opinion was provided by one of the authors (JE) to estimate a relatively low Se (20%, with 95% confidence of being under 33%) based upon the expense and complexity of CKD diagnosis and high Sp (99.5%, with 95% confidence of being above 98%) values (McV Messam et al., 2008). Beta prior distributions were parameterised using the BetaBuster program (Su, 2012).

Risk factor analysis

The nested case-control study used multivariable logistic regression modelling for risk factor analysis. Demographic variables evaluated included individual breed (breeds with ≥12 dogs), purebred/crossbred, KC registered breed status, sex, neuter status, insured status, weight category and age category. Variables with a liberal association in univariable logistic regression modelling (P-value <0.20) were evaluated using multivariable logistic regression modelling. Model building used manual backwards elimination. All eliminated factors were re-evaluated for confounding effects within the provisional final model before confirming their removal. Collinearity between variables considered in the multivariable modelling process was assessed with Spearman's correlation coefficient and Chi-squared tests as appropriate. Biologically meaningful pairwise interactions were assessed between the final model variables. The effect of clustering at the clinic level was evaluated by comparing the final non-random effects multivariable model with a mixed-effects model that included the clinic attended as a random effect (Dohoo et al., 2009). Model fit diagnostics were evaluated using the Hosmer-Lemeshow test (Hosmer & Lemeshow, 2000) and the area under the receiver operating characteristic (ROC) curve.

Survival analysis

Survival times from CKD diagnosis to CKD-related death were estimated using the Kaplan-Meier method with differences between variable categories evaluated by the log-rank test.

Hazard ratios for CKD death were estimated using Cox proportional hazard modelling. Demographic (individual breed [breeds with ≥12 dogs], purebred/crossbred, KC registered breed status, sex, neuter status, insured status, weight category and age category) and blood biochemical (blood urea nitrogen [BUN], IRIS Stage) variables with P<0.20 in univariable Cox proportional hazards regression models were evaluated using multivariable Cox proportional hazard modelling. Model fitting used a manual backwards elimination approach with significance set at P<0.05. The proportionality assumption was tested using Schoenfeld and scaled Schoenfeld residuals and the fit of the final model to the data was evaluated using Cox-Snell residuals.

RESULTS

Practices

During the study period, 89 practices shared EPR data covering 107,214 dogs with 803,132 EPR records with the VetCompass project. These practices were distributed mainly across central and south-east England (Fig. 3).

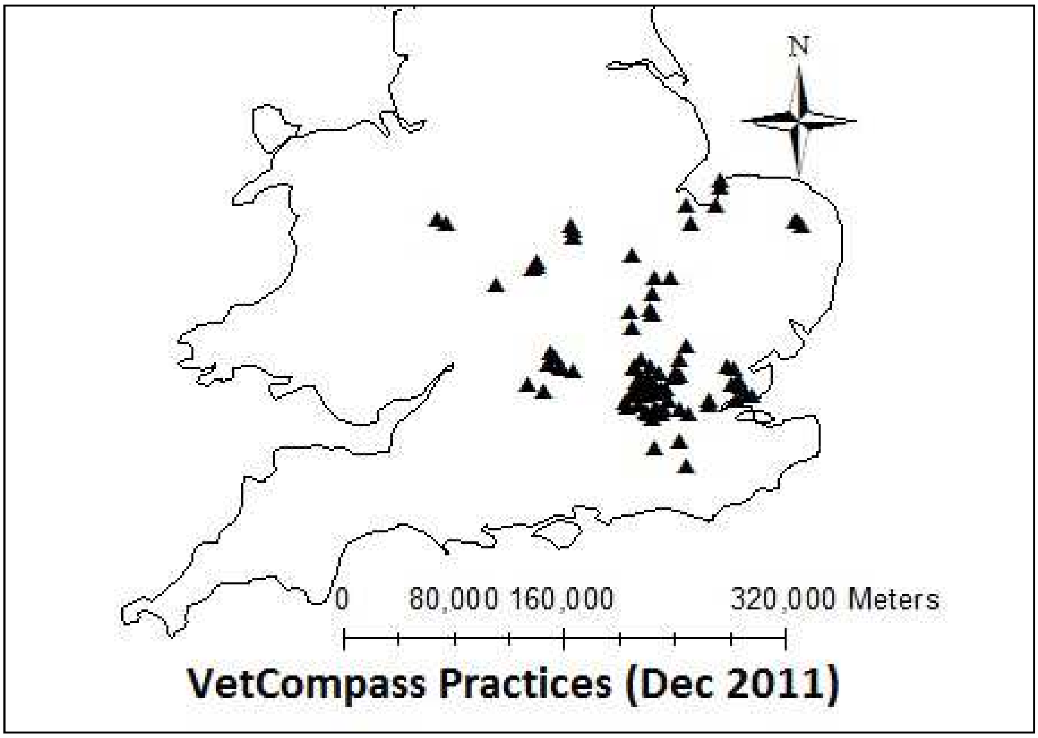


Fig. 3 Map of the southern UK showing the distribution of practices collaborating with the VetCompass project

Prevalence

Overall, 228 dogs met the CKD case definition, giving a CKD AP of 0.21%. Using Bayesian inference, the CKD TP was estimated to be 0.37%.

Risk factor analysis

Multivariable logistic regression modelling results indicated that age group, insurance status and breed were important risk factors for CKD diagnosis. Inclusion of the clinic attended as a random effect within a mixed-effect model did not improve the model. No significant interactions were found between the final model variables. The Hosmer-Lemeshow test indicated good model fit. The area under the ROC curve was 0.8783, indicating good CKD discrimination (Hosmer & Lemeshow, 2000).

Survival analysis

Kaplan-Meier survival curves indicated that the overall median survival time from CKD diagnosis until CKD-related death was 226 days. Blood urea nitrogen concentration and IRIS stage at diagnosis were statistically significantly associated with survival but no significant differences were identified for any demographic variables assessed.

Cox regression modelling indicated that increasing BUN concentration and IRIS stage were associated with an increased hazard of CKD death. There was no evidence of interaction in the final model. The model assumptions were met and the model fitted the data adequately.

DISCUSSION

This large epidemiological study of over 100,000 vet-visiting dogs attending 89 practices is a proof-of-concept for the utility of merged primary practice EPR data for clinical research. The unique insight from primary veterinary practice research derives from collective analysis of the cumulative clinical experience of first-opinion practitioners examining the general population of dogs (Faunt et al., 2007) supplying answers for relevant questions for companion animal health (Holmes & Ramey, 2007). Although paper-based primary-practice studies have been published (Godfrey, 2005; Hill et al., 2006), the substantial manual workload involved has naturally limited their scope. Technological developments in recent years relating to the storage, transfer and analyses of large volumes of electronic data have now reduced practice and patient numbers as a limiting factor for such studies.

Development of systems to collect and analyse anonymised diagnoses from veterinary practices has been identified as critical to fill disorder-prevalence data gaps that constrain current dog breeding reforms (Bateson, 2010). A review of 396 inherited disorders affecting the 50 most common UK dog breeds identified that just 36 conditions had published prevalence estimates and, of these, only 10 were specific to the UK population (Asher et al., 2009). EPR data are generally recorded contemporaneously with the episode of care and by veterinary-trained staff, reducing recall and misclassification bias inherent in questionnaire study designs. Primary practice studies additionally gain from their on-going data collection process that offer a legacy effect with future studies benefitting from ever-increasing practice, patient and temporal data.

Implementation of the VeNom Codes within the PMSs of participating practices is a key component of the VetCompass project. Veterinary practices generally record clinical notes in free-form text format that suffers from the noise introduced by the use of veterinary-specific acronyms, abbreviations, synonyms and common-use codes as well as mis-spellings. Although efforts have been made to automate epidemiologic analyses of free-form text (Lam et al., 2007), the free-form text issue still constrains large scale epidemiological analyses. Natural language programming projects offer potential to overcome this barrier with medical text data (Wang et al., 2010). Allowing clinicians the option to select diagnosis summary terms from a comprehensive but sensibly-sized list is an approach that aims at consistent coding of episodes of care at the time of the event and by the clinician who is best placed to most accurately summarise all available clinical knowledge at that time. The clinician-coded approach can be combined with natural language programming for complementary benefits. The VeNom Code lexicon has been specifically developed to meet the needs of referral and primary practice veterinary surgeons. The Codes are a ‘living list’ controlled by the Venom Coding Group (The VeNom Coding Group, 2012) with regular updates based on clinician feedback. The Codes comprise lists covering species (n=45), dog breeds (767), cat breeds (101), presenting complaints (201), diagnostic tests (39), summary diagnosis terms (2,291) and procedures (780). The term-hierarchical system running behind the VeNom codes greatly facilitates data handling and analysis (e.g. cruciate ligament rupture – cranial → cranial cruciate ligament disorder → cruciate disease finding → stifle disorder finding → joint finding → musculoskeletal finding).

Disorder prevalence studies commonly report apparent prevalence rather than recording true prevalence values because of imperfect clinical tests (Drobatz, 2009). All health evaluations, including anamnesis, clinical examinations and laboratory tests, can be considered as diagnostic tests; combinations of the information gained from these tests then constitute the overall CKD diagnosis test (Dohoo et al., 2009). The apparent prevalence records the level of diagnosed disease rather than the true level of disease. True prevalence can be calculated by formulaic adjustment within the frequentist statistical paradigm if diagnostic sensitivity and specificity are known (Dohoo et al., 2009). However, robust diagnostic sensitivity and specificity information for clinical examinations is rarely available. Bayesian analytic methods are increasingly applied to veterinary epidemiologic data to formally incorporate prior information and expert opinion into prevalence calculations (McV Messam et al., 2008) and to estimate the true prevalence of disease. In the current study, the CKD true prevalence estimation of 0.37% compared with an apparent prevalence estimate of 0.21%, quantifying these perceived limitations of clinician diagnostic sensitivity and specificity.

This study highlights CKD as a disorder with important welfare consequences for dogs. The prevalence shown, although relatively low, is clinically relevant given the unpleasant clinical signs and poor temporal case survival revealed. The identification of demographic and clinical risk factors should facilitate earlier diagnosis, resulting in improved animal welfare.

Restriction of the distribution of practices to central and south-east England may have limited the generalisation from the results to the entire UK. However, the high number of participating practices (89) and veterinary surgeons included should ensure that the clinical protocols leading to diagnoses from participating practices are broadly representative of overall UK practices. Since the time of the reported study, ongoing practice recruitment means that the current distribution of VetCompass practices much more closely reflects the overall UK vet practice map.

In summary, merged datasets of primary practice EPR data are an effective resource for epidemiological analyses. Ongoing practice recruitment and temporal data collection render primary practice EPR projects increasingly effective over time.

ACKNOWLEDGEMENTS

The funding of D. O’Neill by the Royal Society for the Prevention of Cruelty to Animals (RSPCA) for this study is gratefully acknowledged. Idexx Ltd. is also recognised for their support of this study. We thank Peter Dron (Royal Veterinary College) for database development and Marshal Mweu (University of Copenhagen) for OpenBUGS assistance. Special gratitude to the Medivet Veterinary Partnership and their clients who participated in the study.

REFERENCES

APGAW (2009). A healthier future for pedigree dogs. In A healthier future for pedigree dogs: The Associate Parliamentary Group for Animal Welfare, London, UK

Asher, L., Buckland, E., Phylactopoulos, C.L., Whiting, M., Abeyesinghe, S. and Wathes, C. (2011). Estimation of the number and demographics of companion dogs in the UK. BMC Vet. Res. 7, 74

Asher, L., Diesel, G., Summers, J.F., McGreevy, P.D. and Collins, L.M. (2009). Inherited defects in pedigree dogs. Part 1: disorders related to breed standards. Vet. J. 182, 402-11

Bateson, P. (2010). Independent inquiry into dog breeding. Cambridge: University of Cambridge

BBC (2008). Pedigree dogs exposed. Available from http://www.bbc.co.uk/pressoffice/pressreleases/stories/2008/08_august/19/dogs.shtml [Accessed 31 May 2011]

Broom, D.M. (1991). Animal welfare: concepts and measurement. J. Anim. Sci. 69, 4167-75

Brown, S.A. (2007). Management of chronic kidney disease. In BSAVA manual of canine and feline nephrology and urology. edited by Elliott, J. and Grauer, G.F., Quedgeley: British Small Animal Veterinary Association

DiBartola, S.P., Tarr, M.J., Webb, D.M. and Giger, U. (1990). Familial renal amyloidosis in Chinese Shar Pei dogs. J. Am. Vet. Med. Assoc. 197, 483-487

Dohoo, I., Martin, W. and Stryhn, H. (2009). Veterinary Epidemiologic Research. 2nd ed: University of Prince Edward Island.

Drobatz, K.J. (2009). Measures of accuracy and performance of diagnostic tests. J. Vet. Cardiol. 11 (Supplement 1), S33-S40

Faunt, K., Lund, E. and Novak, W. (2007). The power of practice: Harnessing patient outcomes for clinical decision making. Vet. Clin. North Am. Small Anim. Pract. 37, 521-532

Godfrey, D.R. (2005). Osteoarthritis in cats: a retrospective radiological study. J. Small Anim. Pract. 46, 425-429

Haahr, M.,(2012). RANDOM.ORG: True random number service. TSDA. Available from http://www.random.org/ [Accessed 17 March 2012]

Hein, H.E. (1963). Abnormalities and defects in pedigree dogs-II. Hereditary aspects of hip dysplasia. J. Small Anim. Pract. 4, 457-462

Hill, P.B., Lo, A., Eden, C.A., Huntley, S., Morey, V., Ramsey, S., Richardson, C., Smith, D.J., Sutton, C., Taylor, M.D., Thorpe, E., Tidmarsh, R. and Williams, V. (2006). Survey of the prevalence, diagnosis and treatment of dermatological conditions in small animals in general practice. Vet. Rec. 158, 533-9

Hodgman, S.F.J. (1963). Abnormalities and defects in pedigree dogs–I. An investigation into the existence of abnormalities in pedigree dogs in the British Isles. J. Small Anim. Pract. 4, 447-456

Holmes, M.A. and Ramey, D.W. (2007). An introduction to evidence-based veterinary medicine. Vet. Clin. North Am. Equine Pract. 23, 191-200

Hosmer, D.W. and Lemeshow, S. (2000). Assessing the fit of the model. In: Applied Logistic Regression: John Wiley & Sons.

IRIS (2007). International Renal Interest Society (IRIS). Novartis Animal Health Inc.. Available from http://www.iris-kidney.com/index.shtml [Accessed 8 March 2012]

Jones, B.R., Gething, M.A., Badcoe, L.M., Pauli, J.V. and Davies, E. (1989). Familial progressive nephropathy in young Bull Terriers. N. Z. Vet. J. 37, 79-82

Lam, K., Parkin, T., Riggs, C. and Morgan, K. (2007). Use of free text clinical records in identifying syndromes and analysing health data. Vet. Rec. 161, 547-551

Lees, G., Helman, R., Homco, L., Millichamp, N., Hunter, J. and Frey, M. (1998). Early diagnosis of familial nephropathy in English cocker spaniels. J. Am. Anim. Hosp. Assoc. 34, 189-195

Macdougall, D.F., Cook, T., Steward, A.P. and Cattell, V. (1986). Canine chronic renal disease: Prevalence and types of glomerulonephritis in the dog. Kidney Int. 29, 1144-1151

McAloose, D., Casal, M., Patterson, D.F. and Dambach, D.M. (1998). Polycystic kidney and liver disease in two related West Highland White terrier litters. Vet. Pathol. Online 35, 77-80

McGreevy, P.D. and Nicholas, F.W. (1999). Some practical solutions to welfare problems in dog breeding. Anim. Welf. 8, 329-341

McV Messam, L.L., Branscum, A.J., Collins, M.T. and Gardner, I.A. (2008). Frequentist and Bayesian approaches to prevalence estimation using examples from Johne's disease. Animal health research reviews / Conference of Research Workers in Animal Diseases 9, 1-23

Murray, J.K., Browne, W.J., Roberts, M.A., Whitmarsh, A. and Gruffydd-Jones, T.J. (2010). Number and ownership profiles of cats and dogs in the UK. Vet. Rec. 166, 163-168

Nelson, R.W. and Couto, C.G. (2009). Small animal internal medicine. 4th ed. ed. Edinburgh: Mosby.

PFMA (2012). The Pet Food Manufacturers' Association 'Statistics'. The Pet Food Manufacturers' Association. Available from http://www.pfma.org.uk/statistics/ [Accessed 14 September 2012]

Polzin, D.J. (2011). Chronic kidney disease in small animals. Vet. Clin. North Am. Small Anim. Pract. 41, 15-30

Polzin, D.J., Osborne, C.A. and Ross, S. (2005). Chronic kidney disease. In: Textbook of veterinary internal medicine: diseases of the dog and cat, edited by S. J. Ettinger and E. C. Feldman. St. Louis, Miss.; Oxford: Elsevier Saunders

RCVS (2012). RCVS Facts 2012: Part 1. edited by Royal College of Veterinary Surgeons. London: Royal College of Veterinary Surgeons

Robinson, D. and Hooker, H. (2006). The UK veterinary profession in 2006 edited by Royal College of Veterinary Surgeons. London: Institute for Employment Studies

Rooney, N.J. (2009). The welfare of pedigree dogs: Cause for concern. J. Vet. Behav. 4, 180-186

Sosnar, M. (2003). Retrospective study of renal failure in dogs and cats admitted to University of Veterinary and Pharmaceutical Sciences Brno during 1999-2001. Acta Vet. Brno 72, 593-598

Su, C.L., (2012). Bayesian epidemiologic screening techniques: Betabuster. Department of Medicine and Epidemiology, University of California, Davis,. Available from http://www.epi.ucdavis.edu/diagnostictests/betabuster.html#Download [Accessed 11 May 2012]

The Kennel Club (2012). The Kennel Club. Available from http://www.thekennelclub.org.uk/ [Accessed 18 November 2012]

The VeNom Coding Group (2012). VeNom Veterinary Nomenclature. VeNom Coding Group. Available from http://www.venomcoding.org [Accessed 4 May 2012]

Thomas, A. (2009). OpenBUGS. Available from http://www.openbugs.info/w/FrontPage [Accessed 15 May 2012]

Thomas, A., O'Hara, B., Ligges, U. and Sturtz, S. (2006). Making BUGS Open. R News 6, 12-17

Thrusfield, M.V. (1989). Demographic characteristics of the canine and feline populations of the UK in 1986. J. Small Anim. Pract. 30, 76-80

Upjohn, M., Jasani, S., Church, D. and Brodbelt, D. (2008). Establishing an Electronic Patient Record (EPR) in first opinion veterinary practice: challenges to overcome. Poster presented at the Society for Veterinary Epidemiology and Preventive Medicine Annual Conference, Liverpool, 2008. Available from http://www.svepm.org.uk/posters/

VetCompass (2012). VetCompass: Health surveillance for UK companion animals. RVC Electronic Media Unit. Available from http://www.rvc.ac.uk/VetCompass [Accessed 29 October 2012]

Wang, X., Chase, H., Markatou, M., Hripcsak, G. and Friedman, C. (2010). Selecting information in electronic health records for knowledge acquisition. J. Biomed. Inform. 43, 595-601

EVALUATION OF THE SPATIAL AND TEMPORAL DISTRIBUTION AND RISK FACTORS OF THE BLUETONGUE SEROTYPE 1 (BTV-1) EPIDEMIC IN EXTREMADURA (SPAIN) IN SHEEP FROM 2007 TO 2011

A.V. PASCUAL*, B. MARTÍNEZ-LÓPEZ, D.U. PFEIFFER, J.C. MORENO, C. SANZ
AND J.M. SÁNCHEZ-VIZCAÍNO

SUMMARY

Extremadura, which holds the largest number of sheep and the second largest number of cattle in the Spain, has been one of the regions most affected by bluetongue (BT) in Spain. BTV-4 was successfully eradicated a year after its incursion in 2004. However, BTV-1 has been present in the area since 2007, despite intensive vaccination programmes. In this study, ArcMap, SatScan and multilevel models in R language were used in order to evaluate the spatial and temporal distribution of BTV-1 in Extremadura from 2007 to 2011 as well as the risk factors that contributed to the occurrence of the infection in this area during those years. Results showed higher levels of BTV-1 in October and November each year. Spatial clusters were identified, especially in the north of Extremadura as the epidemic developed. The risk factors contributing to BTV-1 occurrence significantly changed between 2007 and 2011.

INTRODUCTION

Bluetongue (BT) is a non-contagious, vector-borne disease of domestic and wild ungulates, including sheep, cattle, goats and deer. It is caused by bluetongue virus (BTV), an RNA virus of genus Orbivirus, which includes 26 serotypes (Maan et al., 2011). Each serotype may be considered as a “different disease” because there is no immunological cross-protection and there is a significant difference in the virulence and clinical presentation of the diverse serotypes in different ruminant species (MacLachlan & Guthrie, 2010). BT is characterised by vascular disorders that are commonly seen in sheep, goats and some deer and they are rarely seen in cattle. Cattle, as well as goats and deer, are thought to act as a reservoir for the virus (López-Olvera et al., 2010).

BT is classified as a notifiable disease by the Office International des Epizooties (OIE 2011) for its potential rapid spread and serious economic consequences in affected countries.

---

*Ana Victoria Pascual, Department of Agriculture and Rural Development, Glenree House, Carbane Ind. Est., Newry, BT35 6EF, Northern Ireland, UK.
Email: ana.pascual@dardni.gov.uk / apaslin@hotmail.com

To date, BTV is only transmitted by vectors of the genus *Culicoides* (Walton, 2004). Thus, the geographical distribution of the disease is confined to areas suitable for *Culicoides* spp., which are characterised by particular climatic conditions (Calvete et al., 2008), explaining why BT has been geographically restricted to latitudes between 40 and 50° north and 35° south (OIE, 2011). It has been hypothesised that climate and land use changes in recent years may have influenced and enhanced further distribution of vector-borne diseases, including BT (MacLachlan & Guthrie, 2010). In fact, BTV has re-emerged in Europe since 1998 at much higher latitudes than usual (MacLachlan & Guthrie, 2010). EU funding for the control and eradication of BT between 2002 and 2011 has amounted to over €92 million, with Spain being the recipient of the largest amount of this funding at over €30 million (European Commission Directorate, 2011).

In Spain, after the BTV epidemics between 1956 and 1960 and 40 years of absence of the disease, BTV-2 was introduced in 2000 (Canals i Rosell, 2006). Later, serotypes 4 and 1 were introduced into the Iberian Peninsula in 2004 and 2007, respectively. Windborne introduction of infected *Culicoides* from North African countries has been hypothesised to be the mechanism of BTV-2, 4 and 1 introduction into Spain (Martínez-López et al., 2009; Allepuz et al., 2010). Serotype 8 was also first diagnosed in 2010, but no more outbreaks were reported subsequently. Since 2004, serotypes 4 and 1 have caused high economic losses due to trade restrictions; reduction of the number of animals and changes in their production management practices; and the cost of surveillance, control and eradication measures. Specifically in Extremadura, the control of serotype 4 was achieved within a year by using live and inactivated vaccines. However, serotype 1, which was introduced into the region in 2007, has not yet been controlled. Currently, Extremadura is the one of the few regions in Spain where serotype 1 is still circulating, despite efforts from the national and regional veterinary authorities and the costly surveillance and vaccination campaigns implemented in recent years. Considering that Extremadura is the main producer of sheep in Spain (nearly 3.5 million animals), the fourth largest national producer of goats (275,000 animals) and the second largest producer of cattle (nearly 800,000 animals) (Díaz-Ambrona et al., 2011), the economic impact of BTV-1 not only on this region but the entire country becomes extremely relevant. Moreover, BTV-1 is more virulent (i.e., with more severe symptoms) in infected sheep and goats than other serotypes; this also has important implications for animal welfare. For example, the cost of the BTV-1 vaccination programme in Spain in 2008 was about €50 million (Ministerio de Medio Ambiente y Medio Rural y Marino, 2008). In 2012, the cost of the Spanish national BTV-1 eradication programme was estimated to be over €6 million, including the cost of vaccination, serological testing and entomological surveillance, of which more than 10% was spent in Extremadura (Ministerio de Agricultura, 2012). Considering that most EU countries are in the process of returning to BTV-free status, subsequent to the European 2006 BTV-8 epidemic and several years of vaccination, continued BTV presence in a region is likely to result in restrictions to animal trade within the EU.

Some of the factors associated with BTV-8 occurrence have been described in previous studies. Specifically, *Culicoides* abundance, animal movements, vaccination coverage, animal density and production type have all been identified as significantly affecting BTV-8 risk (Turner et al., 2012). However, the relative importance of these factors may change from one serotype to another and from region to region; this highlights the importance of conducting analyses to identify important risk factors in the region under study.

The aims of this study were (1) to describe the spatial and temporal patterns of the series of BTV-1 epidemics that affected Extremadura from 2007 to 2011; (2) to identify geographical areas at high risk of BTV-1 occurrence; and (3) to determine the most important risk factors that influence BTV-1 occurrence in the region. The outcomes of this study will be used to provide information for risk-based, more cost-effective surveillance and control strategies with the aim of reducing BTV-1-associated economic and animal welfare problems and, ultimately, to more rapidly achieve BTV-1 eradication in the region.

MATERIALS AND METHODS

Study area

Extremadura, which is divided into two provinces, Caceres in the north and Badajoz in the south, is an Autonomous Community of Western Spain with a total area of 41,634 km^2. The region is bordered by Portugal in the west; Castile and Leon in the north; Castile-La Mancha in the east and Andalusia in the south. It has a Mediterranean climate except in the north, where it is Continental. Extremadura has a large and economically important livestock sector. For example, nearly 89% of the human population live in rural areas, of which approximately 32% have a professional association with livestock production.

Data

Information regarding BTV-1 outbreaks, animal census, farm locations, types of production, species and breeds on farm, number of animals vaccinated on each farm and vaccine type used, *Culicoides* spp. trapped and regional, national and international animal movements for each farm were provided by the Junta de Extremadura (veterinary services of the regional government of Extremadura) for each of the years 2007–2011. Specifically, BTV-1 data consisted of the identification number for each farm (ID), location of the infected farm (UTM), day of the confirmation of the outbreak and the number of infected animals on each farm over the specified time period. An outbreak was defined as a BTV-1-infected sheep farm (including mixed farms with sheep and other species) located in Extremadura where BTV-1 was diagnosed either by ELISA or RT-PCR (OIE, 2009) in the regional lab and confirmed and serotyped in the Veterinary Central Laboratory in Algete (National Reference Laboratory for Bluetongue in Spain) from 2007 to 2011. The date of outbreak occurrence was assumed to be the day of laboratory confirmation.

Information on sheep farms consisted of the location (latitude and longitude of the farm centroid), number of animals on a farm (sheep and other ruminant species), breed of sheep, type of production (meat, dairy or both), management system (extensive or intensive), whether the farms were sentinel or not (not vaccinated/vaccinated and part of the BT National Surveillance Program), animal movements and dates of each move into each of the farms from other areas of Extremadura as well as from other Spanish areas and from other countries. The density of cattle around each sheep farm was also computed because cattle may act as a reservoir for disease transmission. This was done using the kernel function implemented in ArcGIS v. 9.3 to produce a raster map of the cattle density in Extremadura. The density of cattle close to the sheep farm was then extracted (with 1 km^2 cell size and 8 km as the bandwidth) (Sanders et al., 2011).

Information related to the vaccination of BTV-1 in Extremadura was also available and incorporated in this study. Specifically, the day of vaccination, number of animals vaccinated and type of vaccine used (i.e., serotype, monovalent and bivalent vaccines). The percentages of animals vaccinated on farms from December to August were computed because this was assumed to be the time period in which vaccination may protect the animals for BTV infection during the same year (i.e., September to December).

Entomological information on the number of *Culicoides* spp. gathered on the Entomological National Surveillance programme in Spain was available (Ministerio de Medio Ambiente, 2011). Specifically, the number of *Culicoides imicola* captured with Mini-CDC traps and the date of capture was used to produce a *Culicoides* raster maps for each year under study by using inverse distance weight (IDW).

Statistical analyses

After cleaning, data were linked into one table for each year under study using Microsoft Office Excel and Access 2007. An extensive descriptive analysis was performed to explore the information provided, to evaluate its completeness and reliability, and to select best candidate predictors to be included into the statistical multilevel model. Specifically, measures of central tendency and dispersion were computed and epidemic curves, maps, graphs, histograms and tables were produced to explore and describe the information available. R-language 2.15.0 was used to compute all these measures and Microsoft Office Excel 2007 and ArcMap version 9.3.1. (ESRI) were used to produce graphs and maps. SaTScan version 9.0 was used for the statistical spatial analysis of clusters.

A multilevel logistic regression model was used to evaluate the factors that contributed to BTV-1 occurrence in each of the five epidemic waves of BTV-1 in Extremadura (2007 to 2011). The unit of analysis was the farm. We included the municipality as a random effect to account for similarities within municipalities, such as practices of local veterinarians or similarities in environmental/climatic conditions. Up to 71 variables were tested as potential risk factors for BTV-1 occurrence. Specifically, quantitative risk factors were evaluated both in the binary (presence/absence; values above/below the median or using the algorithm of Jenks) (Jenks and Caspall, 1971) and standardised format ({Number-Mean}/Standard Deviation). These 71 variables (including standardised and binary) were: number of bovine, caprine and ovine animals; total number of animals; species in the same farm (bovine/caprine/ovine); breed of sheep (12 different breeds); type of production on each species (dairy/meat/mixed/bullfighting); management type (extensive/intensive); latitude and longitude of each farm; number of *Culicoides imicola*; sentinel (Yes/No); outbreak presence (Yes/No); number of outbreaks; time from outbreak to outbreak in the same farm; vaccination (Yes/No); type of vaccine (monovalent/bivalent and of which serotypes); percentage of vaccinated animals; number of movements and number of animals moved from Extremadura, other national movements or international movements; total of animals moved and total number of movements; and density of bovines within 1 km around the farm. Selection of the risk factors to be included into the best-fit model was achieved by first fitting a non-hierarchical model (no random effects) and using a backward elimination process to remove non-significant variables (p>0.1). Once the model with the best fit for 2007 was selected, the random effect was incorporated into the model. The same process was used for the other years (2008–2011) and variables significant in at least one model were retained in the final models for each year, to allow a comparison of factors that contribute to BTV-1 for the different epidemic waves (i.e., 2007–2011). Confounders were evaluated by looking at significant changes (>20%) on estimates after introducing them into the model. The variance inflation factor (VIF) was used to test for collinearity problems (Fox, 2008). The area under the receiver operating characteristics curve (AUC) was used to quantify the model fit. All statistical analyses were conducted by using R-language v.2.15.0 (The R Foundation for Statistical Computing, 2012).

The final logistic multilevel models used to evaluate the risk factors contributing to BTV-1 occurrence in Extremadura from 2007 to 2011 included 10 fixed effects and municipality as a random effect (Table 1). Specifically, the fixed effects incorporated into the model were the standardised number of cattle animals present on farm; the number of sheep on a farm above (1) or equal to/below (0) the median; the standardised number of goats present on a farm; the presence of the *Talaverana* breed on a farm (yes/no); the production system of a farm being extensive; a high (1) or low (0) number of *C. imicola* on a farm; a high (1) or low (0) proportion of BTV-1-vaccinated sheep on a farm; sheep introduced on a farm from other farms within Extremadura (yes/no); cattle introduced on a farm from other farms within Extremadura (yes/no); and a high (1) or low (0) density of cattle around (<1 km radius) the sheep farm. There were no significant interaction terms or confounders.

RESULTS

Spatio-temporal distribution of BTV-1 in Extremadura from 2007 to 2011

The majority of the BTV-1 outbreaks (Table 1) occurred during 2007 (2991 out of 3281 outbreaks, 91%), most likely because the population was immunologically naïve (i.e., not previously exposed or vaccinated) to this serotype. In subsequent years, the outbreaks decreased greatly in number and were concentrated in the north of Extremadura. For most of the epidemic waves, the BTV-1 index case was identified at the end of September (weeks 38–39) with exception of 2011, when the first outbreak was identified in October. The peaks of the epidemics were reached in November, with the exception of 2008 and 2010, when the peaks were reached in October. Similarly, most of the outbreaks were identified in November each year, except in 2010 when the majority of outbreaks were identified in October.

The spatial distribution of the outbreaks significantly differed from year to year (Fig. 1). In 2007, BTV-1 was extensively distributed in the region, with most of the cases allocated in the south, south-east and north-west of Extremadura. During 2008, the disease was more clustered in two areas: south-west and central north-east. From 2009 to 2010, the majority of the disease was located in the northern part of the region (i.e., Caceres). In 2011, only 8 outbreaks were notified; these were dispersed over Caceres with no apparent clusters. No BTV-1 was notified in Badajoz during the last two years of the study period. This was confirmed by a spatial cluster analysis which showed that there were 3 significant ($p<0.001$) BTV-1 clusters in 2007, 2 in 2008, 1 in 2009, 1 in 2010 and none in 2011 (Fig. 1). The relative risk of infection in these clusters was particularly high in the years 2008, 2009 and 2010. The primary clusters during these three years were all in the north of Caceres with a radius of 34, 98 and 76 km, respectively. The farms within these clusters were 19.31 times in 2008, 53.53 times in 2009 and 167.63 times in 2010 as likely to have BTV-1 as those outside.

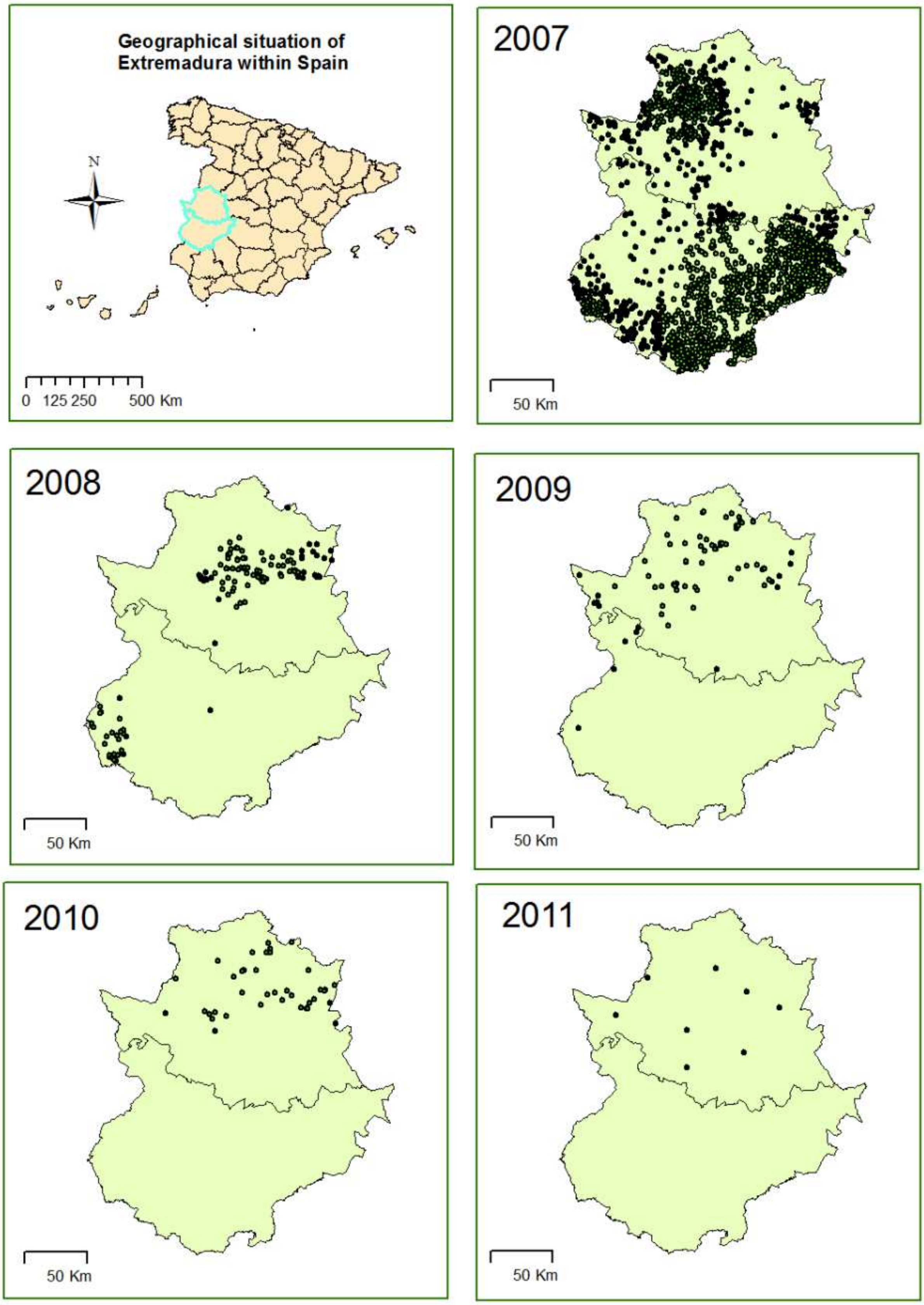


Fig. 1 Spatial distribution and clustered (green dots) BTV-1 outbreaks in Extremadura from 2007 to 2011

Table 2. Number (N) and proportion (%) of BTV-1 outbreaks notified and animals vaccinated by year and species in Extremadura

Year	N Outbreaks Sheep (%)	N Outbreaks Cattle (%)	N Outbreaks Goats (%)	N vaccinated Sheep (%)	N vaccinated Cattle (%)	N vaccinated Goats (%)
2007	2991 (23.35)	10 (0.11)	0 (0)	1,262,566 (29.64)	480 (0.05)	2691 (1.05)
2008	152 (1.27)	2 (0.02)	1 (0)	7,295,172 (100)	1,764,470 (100)	3241 (1.30)
2009	90 (0.84)	0 (0)	2 (0)	4,023,256 (100)	1,261,345 (100)	32 (0.01)
2010	40 (0.39)	2 (0.02)	3 (0)	3,247,070 (100)	984,067 (100)	40 (0.02)
2011	8[a] (0.08)	1 (0.01)	4 (0)	2,205,787 (74.41)	548,612 (57.21)	0 (0)

[a] Includes 2 outbreaks notified in January 2012

Multivariable analysis

The results in Table 2 show that most of the factors studied contributed differently to the risk of BTV-1 occurrence in each of the years. In 2007, significant risk factors for BTV-1 occurrence were having a large number of sheep on farm with an OR of 10.72 and $p<0.0001$, having an extensive management system (OR=2.07, $p<0.0001$), having a large number of *Culicoides* captures on farm (OR=1.81, $p=0.0129$), introducing sheep from other regions of Extremadura (OR=1.32, $p<0.0001$) and having cattle on the farm (OR=1.07, $p=0.03$). In this year, having a high density of cattle around sheep farms and breeding *Talaverana* sheep were significant protecting factors for the disease.

In general, having a large number of sheep on a farm and having cattle on the sheep farms contributed to BTV-1 occurrence in all years, except in 2011. However, other factors, such as having the *Talaverana* breed on a farm, changed drastically from being a protective factor in 2007 (OR=0.46, $p=0.0794$) to becoming a significant risk factor for disease occurrence in 2008 and 2009 (OR=4.73, $p=0.0439$ and OR=3.57, $p=0.0545$, respectively). Farms with more than 90% of the sheep vaccinated tended to be at lower risk of disease occurrence, since ORs were below one throughout the study period; however, there was no statistical evidence ($p>0.05$). Conversely, there was very strong statistical evidence that purchasing sheep from other areas of Extremadura in 2007 was a risk factor (OR=1.32, $p<0.0001$). There was also some evidence of this in 2010 (OR 2.69, $p=0.0578$). For the rest of the years, the ORs were above one, although no statistical evidence was found. Results for farms that purchased cattle from local areas in Extremadura did not give statistical evidence of being a risk factor for BTV-1 occurrence until 2011, when results suggested that these farms had 9.9 times the odds of being infected with BTV-1 than farms that did not purchase cattle from other Extremadura areas (p=0.0779). Similarly, during 2007, farms situated in an area with a high density of cattle within a 1 km radius had 0.74 times the odds of being infected than farms with a lower density of cattle within a 1 km radius of the farm (p=0.0183). The results were similar in 2008, but the pattern changed for 2009, 2010 and 2011, showing an OR above 1, although it was not significant.

Table 2. Odds ratios (ORs) and confidence intervals (CI) obtained after fitting logistic multilevel models for each year from 2007 to 2011. Variance of the random effect, municipality, for each year under study

	2007 OR (95% CI)	2008 OR (95% CI)	2009 OR (95% CI)	2010 OR (95% CI)	2011 OR (95% CI)
Cattle (standardised)	1.07[a] (1.01-1.13)	1.16[b] (1.06-1.27)	1.03 (0.85-1.26)	1.36[b] (1.09-1.70)	1.17 (0.95-1.44)
N. Sheep (>median)	10.72[b] (9.23-12.45)	22.96[b] (10.10-52.18)	7.27[b] (3.25-16.29)	3.47[a] (1.20-10.00)	1.72 (0.09-31.15)
N. Goats (standardised)	0.97 (0.91-1.04)	1.03 (0.88-1.21)	0.8 (0.45-1.45)	0.98 (0.75-1.28)	0 (0-Inf)
Talaverana (Y/N)	0.46[a] (0.19-1.09)	4.73[a] (1.04-21.45)	3.57[a] (0.98-13.09)	0.93 (0.21-4.13)	0 (0-Inf)
Extensive (Y/N)	2.07[b] (1.49-2.89)	7.24 (0.67-78.62)	2.75 (0.25-30.58)	0.62 (0.09-4.40)	0.18 (0.01-2.84)
N. *Culicoides* (>1939)	1.81[a] (1.13-2.89)	1.88 (0.61-5.84)	0.97 (0.25-3.69)	0.47 (0.07-3.34)	0 (0-Inf)
Vaccine (>90% coverage)	0.35 (0.02-5.08)	0.8 (0.42-1.52)	0.72 (0.34-1.53)	1.34 (0.54-3.33)	0 (0-Inf)
Sheep moved in (Y/N)	1.32[b] (1.15-1.51)	2.02 (0.61-1.69)	1.45 (0.75-2.82)	2.69[a] (0.97-7.49)	1.8 (0.14-23.43)
Cattle moved in (Y/N)	0.83 (0.64-1.06)	1.24 (0.70-2.19)	1.1 (0.50-2.40)	0.47 (0.11-1.99)	9.9[a] (0.77-126.60)
Cattle density near sheep farms (>37)	0.7[a] (0.58-0.95)	0.78 (0.40-1.51)	1.85 (0.73-4.66)	1.22 (0.36-4.15)	1.88 (0.14-26.30)
σ^2 Municipality	4.85	39.21	17.515	60.782	1.11e-08

[a] Strong evidence that the variable contributes to BTV-1 occurrence

[b] Very strong evidence that the variable contributes to BTV-1 occurrence

The large values obtained for the AUC of the ROC curves (0.91 in 2007, 0.96 in 2008, 0.98 in 2009, 0.99 in 2010 and 0.93 in 2011) reflected a good fit of the model for all of the years. The relative risks of infection for each farm obtained by our models had a very high correlation with the real risk for the farms (infected/not infected in reality). Visual assessment of the risk maps of BTV-1 occurrence also suggested good agreement between predicted risk and reported outbreak areas. The risk was closer to one in areas where the outbreaks occurred and very close to zero where there were no reported outbreaks. The residuals were also very close to zero, indicating good model fit. However, it is of note that, with the exception of 2007, residuals had a tendency to be negative; this suggests an overestimation of BTV-1 risk.

DISCUSSION

This is the first in depth study to describe and evaluate the risk factors contributing to the different epidemic waves of BTV-1 from 2007 to 2011 in one of the most BTV-affected regions of Spain. Some of our results are in agreement with previous research on other BT serotypes, such as the importance of Culicoides distribution/abundance in the region (Allepuz et al., 2010). However, novel and interesting results were obtained on some variables, such as the presence of cattle on sheep farms, cattle density around the sheep farms, vaccination coverage on a farm, incoming sheep/cattle movements and local sheep breeds that specifically apply to Extremadura; these factors may guide decision makers in future interventions in the region.

The temporal description shows that in alternative years the peak of outbreaks are in either October (2008 & 2010) or November (2007, 2009 & 2011), particularly in 2011 the majority of outbreaks were confirmed late in the year (November & December 2011 and January 2012). This could be because of increased awareness of BTV from farmers and veterinarians after a year with late confirmation of outbreaks in the previous year. Further study on climatic conditions for each year could also explain this phenomenon.

Spatial clustering in the northern parts of Extremadura could be explained by a combination of significant variables found in the statistical study. More specifically, these clusters coincide with higher cattle density areas and with areas where the *Talaverana* breed is located; both variables were strongly related to disease occurrence in some years of the study.

Changes in some of the risk factors from 2007 to 2011 suggested that interventions should change, depending on the status/evolution (i.e., control vs. eradication) of the BTV-1 epidemic. For example, at the beginning of the epidemic, in 2007, when most of the sheep population was naïve, factors such as sheep density, sheep movements, *Culicoides* abundance and management practices were very important risk factors for BTV-1 occurrence. In contrast, in 2009 and 2011, when outbreaks significantly decreased thanks to natural immunisation of the sheep population as well as the vaccination campaigns, the role of reservoirs, such as cattle, seems to be crucial for maintaining the disease in the region. In fact, the presence of cattle on a farm or in close proximity (<1 km radius), as well as the purchase of cattle, seems to increase the risk of BTV-1 on sheep farms. This may indicate that the presence of cattle close to sheep flocks may contribute to BTV-1 persistence of the disease in the northern areas of Extremadura. These findings agreed with other previous studies (Brewer & Maclachlan, 1992), which highlighted that cattle were asymptomatic and long-lasting viraemic reservoirs of the virus and may indicate the importance of targeting vaccination in those cattle farms near sheep farms, particularly in eradication phases.

The results also suggested that large sheep farms and farms purchasing sheep were at higher risk of BTV-1 occurrence. Similarly, high numbers of *Culicoides* captured on a farm as well as extensive management practices were also significant risk factors for BTV-1 occurrence, mainly in the initial stages of the epidemic. Certainly, animals in extensive, large sheep farms may be more exposed to *Culicoides* bites than intensive, small farms

(Meiswinkel et al., 2000). However, these factors decreased in importance in 2010 and 2011, which again, may suggest that other risk factors become more important for disease occurrence in the final phases of the eradication program. The contradictory results found for *Talaverana* sheep could be explained, at least in part, by considering the clustered spatial location of this breed. In fact, the area where the *Talaverana* breed was located was not affected at the beginning of the epidemic (i.e., 2007), but later was; this may indicate why results change from being protective to a risk factor. It would be interesting to conduct further studies on susceptibility or management of this breed because no other breeds were positively or negatively associated with BTV-1 occurrence. Differences in breed susceptibility have been reported previously (Ward, 1994). Regarding vaccination, 90% or more vaccinated sheep on a farm clearly tends to be a protective factor over the years and, together with other control measures, may explain the drastic decrease in BTV-1 outbreaks throughout the study period. However, to properly and more deeply evaluate vaccine efficacy in the region, it should be considered that vaccination strongly depends on the administration process/routine on a farm. It should be also taken into account that not all susceptible animals were vaccinated (not compulsory in goats or red deer), that many types of vaccine were used (different trademarks, either mono- or bivalent) and that vaccination in cattle and sheep became optional after June 2011. Thus, it will be interesting to conduct future studies to properly assess the efficacy of vaccination in the field and to evaluate what may be the most cost-effective vaccination strategy.

Another aspect that should be considered is that data were collected across five different years and, therefore, changes in the collection process during that period may have incorporated some bias into study. Moreover, cattle and goat BTV-1-infected premises were intentionally excluded from the study to avoid bias because under-reporting of cases were expected for these farms, since clinical symptoms are usually not observed. Some under-reporting bias may be expected for some sheep farms, which may have some impact on the results presented herein. However, under-reporting should be limited because farmers usually notify diseased animals to receive economic compensation from the government.

Another assumption was that, for *Culicoides* abundance, the maximum number of *C. imicola* captured in the trap from August to November each year was used. This was based on previous studies (Acevedo et al., 2010) and on factors such as the amount of *Culicoides* spp. per season collected in this region, the DEI (days extrinsic incubation), the lifespan of the mosquito, the average incubation period and the mean length of viraemia in sheep and cattle (OIE, 2009). Moreover, data for *C. imicola* was sparse for the last two years of the study (2010–2011). Therefore, pooled data from 2007 to 2009 were also used in the analysis for this period to avoid bias. All of this could have diluted the results towards the no effect of *C. imicola* on BTV-1 infection in those last years. It was also assumed that *Culicoides* spp. do not fly more than a few hundred metres from where they breed and feed (Calvete et al., 2008; Sanders et al., 2011). This was taken into account when selecting the 1 km radius to calculate the density of cattle, which may act as a reservoir, around sheep farms. Because of some missing geographical coordinates (561 farms out of 55,406, approximately 1% of the farms) for some farms, the centroid of the municipality where the farms were located was used. Some bias is likely to occur in all maps presented, as well as in data on *C. imicola* and the density of cattle, leading to an under- or an overestimation of the risk. Moreover, the Jenks algorithm was used to categorise the density of cattle and *Culicoides* as high or low; this was considered reasonable based on the distribution of available data. Due to a lack of reliable information on wildlife BTV reservoirs, the presence/abundance of red deer or other wild ruminant species in the region, which may act as reservoirs, was not considered. It will be interesting for future studies to evaluate the role that wildlife and other aspects, such as environmental/climatic factors, may have in the persistence of BTV-1 in the north of Extremadura. Also, it would be very interesting to compare qualitatively and quantitatively the BTV-1 epidemic versus the BTV-4 epidemic (2004–2005). BTV-4 was successfully eradicated in the same area under similar conditions. A comparative study between BTV-1 and BTV-4 in Extremadura could give more information on what aspects may contribute to faster and most effective control of BTV in the region.

Based on these results and to achieve BTV-1 eradication and control, it would be recommended to continue vaccination on high-risk premises, particularly on large, mixed (with cattle) sheep farms and on cattle farms close to sheep farms. Also, farms that purchased sheep and/or cattle from other areas in Extremadura should be targeted for vaccination and surveillance. From the available data, the farms with these characteristics in Extremadura in 2011, where intervention would be recommended, sum to a total of 708 farms (out of a total of 17,045): 114 sheep farms and 594 cattle farms.

ACKNOWLEDGEMENTS

We would like to thank to the staff at the Local Veterinary Services in Junta de Extremadura for collecting and making the data available for this study. Beatriz Martínez-López holds a Juan de la Cierva post-doctoral contract (JCI-2011-10724). We also thank the SUAT-VISAVET team at the Universidad Complutense of Madrid, for their support and assistance.

REFERENCES

Acevedo, P., Ruiz-Fons, F., Estrada, R., Márquez, A.L., Miranda, M.A., Gortázar, C. and Lucientes, J. (2010). A broad assessment of factors determining *Culicoides imicola* abundance: Modelling the present and forecasting its future in climate change scenarios. PLoS ONE 5, e14236

Allepuz, A., García-Bocanegra, I., Napp, I., Casal, S., Arenas, J., Saez, A. and González, M.A. (2010). Monitoring bluetongue disease (BTV-1) epidemic in southern Spain during 2007. Prev. Vet. Med. 96, 263–271

Brewer, A.W. and Maclachlan, N. J. (1992). Ultrastructural characterization of the interaction of bluetongue virus with bovine erythrocytes in vitro. Vet. Pathol. 29, 356–359

Calvete, C., Estrada, R., Miranda, M.A., Borrás, D., Calvo, J.H. and Lucientes, J. (2008). Modelling the distributions and spatial coincidence of bluetongue vectors *Culicoides imicola* and the *Culicoides obsoletus* group throughout the Iberian peninsula. Med. Vet. Entomol. 22, 124–34

Canals i Rosell, A. (2006). Lengua azul o fiebre catarral ovina [Online]. http://www.produccion-animal.com.ar/sanidad_intoxicaciones_metabolicos/infecciosas/ovinos/02-lengua_azul.pdf: Veterinari del Departament de Salut, Generalitat Catalunya

Díaz-Ambrona, M., Seoane, A., Spiegelberg, P., Rábade Rodriguez, M.T. and Vázquez De Prada Grande, M.A. (2011). Resultado de la Encuesta Nacional de ganado Ovino-Caprino. Ministerio de Agricultura, Alimentación y Medio Ambiente Secretaría General Técnica

European Commission Directorate (2011). Report on the outcome of the EU co-financed animal disease eradication and monitoring programmes in the MS and the EU as a whole. F. C. E. C., Rue de la Loi 200 1049, Brussels

Fox, J. (2008). Applied regression analysis and generalized linear models, Second Edition. Sage, Thousand Oaks

Jenks, F. and Caspall, F.C. (1971). Error on choroplethic maps. Definition, measurement, reduction. Ann. Assoc. Am. Geogr., 61, 217–244

López-Olvera, J.R., Falconi, C., Férnandez-Pacheco, P., Fernández-Pinero, J., Sánchez, M.A., Palma, A., Herruzo, I., Vicente, J., Jiménez-Clavero, M.A., Arias, M., Sánchez-Vizcaíno, J.M. and Gortázar C. (2010). Experimental infection of European red deer (*Cervus elaphus*) with bluetongue virus serotypes 1 and 8. Vet. Microbiol. 145, 148–152

Maan, S., Maan, N.S., Nomikou, K., Batten, C., Antony, F., Belaganahalli, M.N., Samy, A.M., Reda, A.A., Al-Rashid, S.A., El Batel, M., Oura, C.A. and Mertens, P.P. (2011). Novel bluetongue virus serotype from Kuwait. Emerging Infect. Dis. 17, 886–889

MacLachlan, N.J. and Guthrie, N.J. (2010). Re-emergence of Bluetongue, African horse sickness and other Orbivirus diseases. Vet. Res. 41, 35

Martínez-López, B., Perez, A.M. and Sánchez-Vizcaíno, J.M. (2009). Identifying equine premises at high risk of introduction of vector-borne diseases using geo-statistical and space-time analyses. Prev. Vet. Med. 100, 100–108

Meiswinkel, R., Baylis, M. and Labuschagne, K. (2000). Stabling and the protection of horses from *Culicoides bolitinos* (Diptera: Ceratopogonidae), a recently identified vector of African horse sickness. Bull. Entomol. Res. 90, 509–515

Ministerio de Agricultura, Alimentacion y Medio Ambiente (2012). Programa de Lengua Azul España 2012. In: AGRARIA, D. G. D. S. D. L. P. (Ed.). http://rasve.mapa.es/publica/programas/NORMATIVA%20Y%20PROGRAMAS%5CPROGRAMAS%5C2012%5CLENGUA%20AZUL%5CPROGRAMA%20LENGUA%20AZUL%202012.PDF.

Ministerio de Medio Ambiente y Medio Rural y Marino (2008). Evolución Lengua Azul en España durante el año 2008. RASVE (Ed.) http://rasve.mapa.es/Publica/Sanidad/documentos/Informe_LA_RASVE_2008.pdf.

Ministerio de Medio Ambiente, Medio Rural y Marino (2011). Survey programme for Bluetongue. Approved* for 2012 by Commission Decision 2011/807/EU. Spain. In: Health & Consumers Directorate-General, U. G.-V. P. & SANCO/10817/2012 (Eds.). European Commission.

OIE (2009). Bluetongue and epizootic haemorrhagic disease. In: 2009, Manual of Diagnostic Tests and Vaccines for Terrestrial Animals.Vol1, Chapter 2.1.3. Paris. http://web.oie.int/eng/normes/MMANUAL/A_Index.htm

OIE (2011). Bluetongue. Aetiology epidemiology diagnosis prevention and control references. OIE Technical Disease Cards. http://www.oie.int/fileadmin/Home/eng/Animal_Health_in_the_World/docs/pdf/BLUETONGUE_FINAL.pdf

Sanders, C.J., Shortall, C.R., Gubbins, S., Burgin, L., Gloster, J., Harrington, R., Reynolds, D.R., Mellor, P.S. and Carpenter, S. (2011). Influence of season and meteorological parameters on flight activity of *Culicoides* biting midges. J. Appl. Ecol. 48, 1355–1364

Turner, J., Bower, R.G. and Baylis, M. (2012). Modelling bluetongue virus transmission between farms using animal and vector movements. Sci. Rep. 2, 319

Walton, T.E. (2004). The history of bluetongue and a current global overview. Vet. Ital. 40, 31–38

Ward, M.P. (1994). The epidemiology of bluetongue virus in Australia - a review. Aust. Vet. J. 71, 3–10

FARMERS' PERCEPTIONS AND DISEASE CONTROL

FARMERS' PERCEPTIONS, ATTITUDES, INFLUENCES AND MANAGEMENT OF INFORMATION IN THE DECISION-MAKING PROCESS FOR DISEASE CONTROL

P. ALARCON[*], B. WIELAND, A. MATEUS AND C. DEWBERRY

SUMMARY

Twenty English pig farmers were interviewed and asked to describe the decision-making process for two diseases they had experienced. In addition farmers' opinions on key information sources for disease control and on the research conducted on risk factors for post-weaning multi-systemic wasting syndrome (PMWS) were investigated. A qualitative analysis of the transcribed interview was done through the use of text templates.

Drivers for disease control were 'mortality', 'feeling of entering in an economically critical situation', 'animal welfare' and 'feeling of despair'. In some non-sustainable situations, other producers seemed to influence considerably farmers' decision-making process. 'Lack of knowledge', 'farm structure and management barriers' and 'economic constrains' were identified as behavioural barriers. 'Lack of communication', 'not knowing where to look' and 'information bias' were themes associated with research from academia. Discussion of PMWS research findings identified positive and negative themes. Overall, insights gained are important for developing communication strategies for disease control.

INTRODUCTION

Diseases can have a substantial negative impact on farms. Accomplishing effective communication of knowledge is essential to optimize farmers' decision-making process for disease control and, therefore, to minimize its impact. However, some studies have shown that, despite the onset of major knowledge transfer programs, effective communication was not achieved (Iles, 2003; Noremark et al., 2009). This indicates that farmers' perceptions and factors affecting behaviour need to be better understood for designing effective knowledge transfer strategies. Nowadays, the amount of information and number of information sources available to farmers is increasing and the resulting information management presents an enormous challenge. Farmers' perceptions and attitudes towards different information sources can significantly impact the way information is managed and decisions are made. An important part of this information is the one derived from research. In the United Kingdom, the Department of Food and Rural Affairs' (DEFRA) budget for evidence-based research on animal health and welfare was £63.2 million for the year 2011/12 (Anonymous, 2011). In the European Union a total of €1,935 million was budgeted on food, agriculture and fisheries

---

[*] Pablo Alarcon, Royal Veterinary College, University of London, London, AL9 7TA, UK. Email: palarcon@rvc.ac.uk

research for the period 2007-2013 (Anonymous, 2007). These amounts do not account for all the private investments in research in the farming industry. To ensure that research findings reach effectively the producers and support them in their decision-making process it is essential not only to optimize disease control and increase competitiveness, but also to ensure that research efforts and outputs have a real impact.

Only few recent studies have investigated farmers' motivators and decision-making process for disease control, and these mainly focussed on the dairy and beef industry (Gunn et al., 2008; Heffernan et al., 2008; Ellis-Iversen et al., 2010). Many of the socio-psychological studies done in the farming sector have used or adapted the Theory of Reason Action (TRA) or the Theory of Planned Behaviour (TPB) for the investigation of farmers' behaviours (Garforth et al., 2004; Gunn et al., 2008; Ellis-Iversen et al., 2010). These theories suggest that actual behaviour may be predicted by the strength of the intention in conducting the behaviour (Fig. 1). The strength of 'behavioural intention' depends on a combination of person's attitudes and social norms. Attitudes refer to the person's positive or negative belief on the effectiveness of the behaviour ('outcome belief') and on its belief on the consequence of this behaviour ('outcome evaluation'). Subjective norms reflect the person's perceptions on the social pressure to perform the behaviour ('subjective beliefs') and the person's motivation to comply with these norms ('motivation to/not to comply'). However, TPB introduces a new element called 'Perceived behavioural control' (PBC), which accounts for the individual's belief in being able to achieve the behaviour ('control belief or self-efficacy') and also for the factors perceived to difficult achieving the behaviour ('power of control'). Ajzen (1991) hypothesized that PBC not only affects intention, but it is also directly related to actual behaviour. These theories have been proven effective in predicting and explaining behaviours (Armitage & Conner, 2001; Jackson et al., 2006).

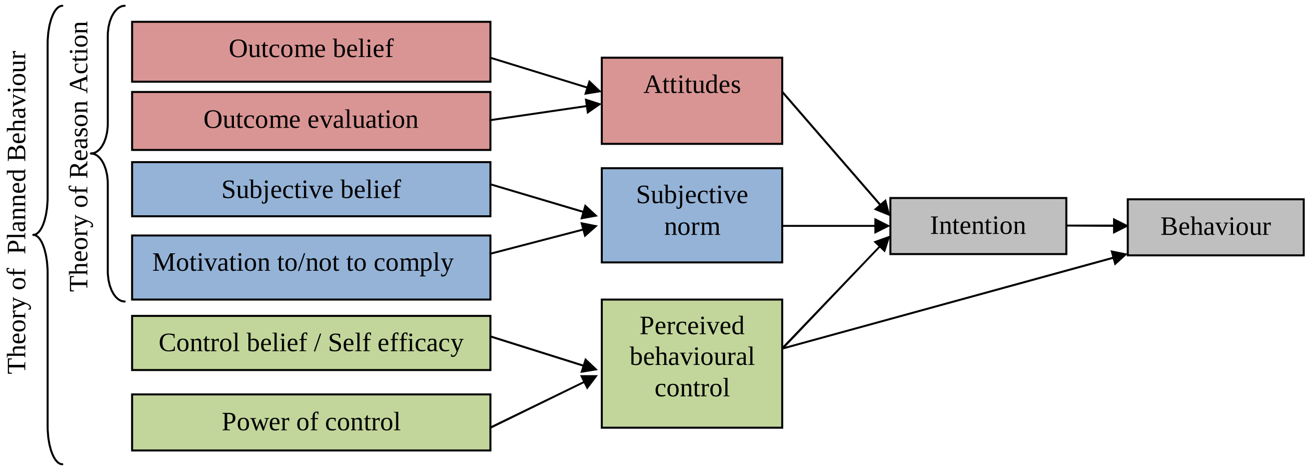


Fig. 1 Theory of planned behaviour (adapted from Garforth et al. (2004) and Armitage & Conner (2001))

The objectives of this study were (1) to explore the factors involved in pig farmers' decision-making process for the control of complicated diseases or ill-defined/structured disease situations; (2) to investigate farmer's attitudes and perception towards different information sources; and (3) to assess farmers' opinions on the research conducted on post-weaning multi-systemic wasting syndrome (PMWS) severity risk factors (Alarcon et al., 2011)

MATERIALS AND METHODS

Data collection

A study involving 20 English pig farmers was conducted between June and July 2011. Eighteen farmers were selected from the PCV2 vaccination programme conducted by BPEX, the English pig levy payer association. Two farmers were recruited through staff at the Royal Veterinary College. Data were collected through face-to-face interviews with farmers using a combination of a semi-structured questionnaire and closed questions. Interviews with farmers were recorded and then transcribed for qualitative analysis purposes. A thematic analysis of the text was done through the use of templates, as described by King (2004).

Qualitative research

Interview method: For this study an initial interview template of themes was created. It was composed of three parts (higher degree nodes),which were related to (1) decision making process for disease control, (2) perceptions and attitudes towards sources of information, and (3) perceptions on PMWS research findings.

The first part aimed at understanding the factors involved in farmers' decision-making process for disease control, with emphasis on the management of information sources. Farmers were asked to identify two diseases that they had experienced in their farms and which were difficult to recognize and/or to control. These disease situations were classified as 'ill-defined' or 'ill-structured' disease problems, as no straightforward cause or solution was known or currently available. These situations also involved important decisions to be made by farmers, who therefore needed to optimize the management of information and its sources. In this context, and to facilitate and aid farmers' narrative of their experience, the questionnaire focussed on three important steps; (1) recognizing a disease problem, (2) deciding the need to control, and (3) deciding how to control it (Second order nodes). For the last step, in order to classify the factors identified in a meaningful framework, a structure related to the theory of planned behaviour was used (Ajzen, 1991). Factors were therefore classified as related to (1) attitude and beliefs, (2) subjective norms, or (3) perceived behavioural control (third order nodes).

The second part of the interview aimed at exploring farmers' perceptions and attitudes towards five key sources of information for disease control, including research from academia. It also aimed at identifying general factors related to barriers to knowledge ('general problems') using two different questions: 1) 'What were the difficulties to get an understanding of the disease issues you have experienced?', and 2) 'What are the main issues concerning the availability of information on pig health problems?'. However, factors related to general information problems obtained throughout the interview were also classified in this part of the template.

The last part of the interview template aimed at exploring farmers' perceptions of the results of a study investigating PMWS severity risk factors. For this purpose farmers were provided with the scientific report (Alarcon et al., 2011). The study and its findings were explained orally by the researcher. After the explanation, farmers were asked to provide their views on the study.

Qualitative analysis: In the initial interview template, a number of themes were defined *a priori*, which corresponded to the areas of investigation of this study. Using this template as a framework, salient farmers' opinions, perceptions, feelings, concepts and ideas were grouped as themes and classified in the template according to their relationship with the existing codes and through careful reading of the interview text. It is important to note that themes could represent a group of observations done by several farmers or be composed of just one observation from one particular participant. A hierarchical coding structure was used in order to allow various degrees of specificity in the analysis. Higher order nodes represent broad, general themes, while lower level codes represent more specific themes. The template structure was modified according to the findings obtained from the interviews and new codes were added when deemed necessary. For analysis purposes, all farmers were treated anonymously. Qualitative analysis was performed using NVivo 8 (QSR international Pty Ltd.)

Quantitative research

A questionnaire composed of closed questions assessed quantitatively the perception of usefulness of different information sources in relation to information provided on diseases and the importance of different type of motivators for disease control. Four types of motivators for disease control were considered in the questionnaire: (1) financial, (2) image and reputation of the farm, (3) sense of pride as a good manager, and (4) animal welfare. These motivators correspond to intrinsic, social, expressive and intrinsic motivators as theorised by Gasson (1973). The sequence of the questions was carefully chosen in order to avoid leading responses. No financial incentives were given to farmers to participate in this study, but personal communication of results from the PMWS severity and risk factor study was given.

RESULTS

The duration of the interviews ranged from 35 minutes to 1 hour and 25 minutes and a substantial amount of data for the different parts of the template was collected. Only in the part covering 'recognition of a disease problem' limited richness was obtained, and therefore this second higher degree node and its inferior codes were removed from the template. Due to the length of the final template, and in order to avoid a superficial analysis of the results, only a selection of nodes is described here. Final template results can be found in Fig. 2.

1. **EXPERIENCE WITH COMPLICATED DISEASES (ill-defined / ill-structured problem)**

1. **Deciding the need to control**
    - A. Motivations for control
        - I. Mortality of pigs
        - II. Low production performance (growth rates, feed conversion rates and fertility problems)
        - III. Animal welfare
        - IV. Economic
            - a. Situation not sustainable
            - b. Economic losses
            - c. Fear of losing contract
            - d. Economically effective control measure available
        - V. Existence of effective control measures
        - VI. Fear of getting the disease
        - VII. Fear of within-farm disease spread
        - VIII. Feeling of despair
        - IX. Reduce abattoir lesions scores
        - X. Reputation of the farm
        - XI. Social responsibility
        - XII. Frequency and persistence of disease problem
    - B. External influence
        - I. Pressure from contractor (abattoir or finishing site)
        - II. Vet advice
        - III. Other farmers (Discussion groups, Personal level)
        - IV. Salmonella reports (government pressure)
        - V. Feed representative and/or nutritionist
        - VI. Family /Partner

2. **Deciding which control measure to use**
    - A. Attitude toward measure
        - I. Outcome belief
            - a. Belief effective
                - i. Because other farmers use it
                - ii. Common sense
                - iii. Logic of good husbandry practices
                - iv. Drastic measures
                - v. Experience on previous use
                - vi. Common practice abroad (abroad is better)
                - vii. Associations observed on-farm
                - viii. Vet confidence on the measure
            - b. Uncertain
                - i. Lack of understanding
                - ii. Not identifying the cause
                - iii. Novel disease
                - iv. Complex disease
            - c. Principles
                - i. Antibiotics used as last resort
                - ii. Limiting injections to pigs
                - iii. Preventive measures
                - iv. Needs to be economically effective

        - II. Outcome evaluation
            - a. Drop in Mortality
            - b. Increase in Growth rates
            - c. Improve fertility
            - d. Reduction in BPHS scores
            - e. Evaluation of the incidence of diseased pigs (clinical signs)

    - B. Subjective norm
        - I. Subjective belief
            - a. Veterinarians
                - i. Own farmer' veterinarian advice
                - ii. Veterinarian at producer conference
                - iii. Visiting veterinarian

Fig. 2 Final template

- b. Other producers
  - i. at pig discussion groups
  - ii. at personal level
  - iii. Through press/magazines
- c. BPEX
- d. Feed representative and/or nutritionist
- e. Breeding company
- f. Pharmaceutical companies and Veterinary colleges (new vaccine development)
- g. Press and magazines

II. Motivation to/not to comply

- a. Positive externalities of control measure
- b. Performing a trial to assess effects
- c. Common sense of the control measure
  - i. Overall good practice code
  - ii. Understanding cause of disease
- d. Trust in veterinarian
- e. Common practice abroad
- f. Cost and Economic effectiveness of control measure
- g. Labour time involved
- h. Feeling of despair
- i. There is no other option
  - i. Other options too expensive
  - ii. Other options not feasible
  - iii. Failure of other measures
  - iv. Don’t know any other control option
- j. Feasibility within the system
- k. Other farmers positive experience and similarity with other farmers disease situation
- l. Veterinarian positive experience
- m. Limiting injections to pigs
- n. Fear of other disease

C. Perceived behavioural control

- I. Control belief
  - a. Conducting trial on-farm
  - b. Coordinated team of different sources
  - c. Veterinarian supervision
  - d. Having a measureable outcome

- II. Power to control
  - a. Economics
    - i. Lack of labour
    - ii. Cost of control measure
    - iii. Lack of cash flow (investment capabilities)
  - b. Farm structure and management
    - i. Strict bio-security facilitates control
    - ii. Limited bio-security on outdoor farms
    - iii. Operating or not at full building capacity
    - iv. Current status of buildings and pens
    - v. Good husbandry facilitates control
  - c. Lack of knowledge
    - i. Not knowing the cause (novel disease)
    - ii. Complexity of disease
    - iii. No control measure available
    - iv. Disease cycle

3. <u>FACTORS INFLUENCING THE USE OF INFORMATION SOURCES</u>

1. General problems

A. Lack of effective communication

- I. Breeding companies not disclosing disease problems
- II. Confidentiality of research
- III. Farmers not sharing information
- IV. Feeling of isolation

Fig. 2 Final template (continued)

- V. No system on disease alert
- B. Too much information / get lost
- C. Poor communication formats
  - I. Confusion of websites
  - II. Difficult to understand/read
- D. Lack of information
  - I. Due to the fact that some are novel disease
  - II. No feedback or solutions proposed with diagnostics/reports
  - III. Lack of international information
- E. Information bias
- F. Lack of time (overload with work)
- G. Lack of knowledge
  - I. Not knowing where to look
  - II. Not having enough knowledge on disease
  - III. Complex disease are difficult to understand
- H. No problem
- I. Relying on very few sources

2. Research / Universities

- B. Positive / advantages
  - I. Becoming more practical / business oriented
  - II. Important for disease control
- C. Negative / Problems
  - I. Lack of funding
  - II. Real research is done abroad
  - III. Belief research should not be funded by farmers or taxpayers
  - IV. Information bias
    - a. Conditions in research do not match pig farm reality
    - b. Low sample size of trials
    - c. Research bias by pharmaceutical companies
  - V. Lack of field expertise by researchers
  - VI. Lack of communication
  - VII. Don't know where to look
  - VIII. Historical
  - IX. Poor communication format
  - X. Not accessible
  - XI. Poor quality of some research done (sample size issue)
  - XII. Some research is not link to their problems
- B. General methods
  - I. Not used
  - II. Informed through BPEX
  - III. Only for vaccine efficacy or equipment information
  - IV. Through workshops
- B. Communication preferences
  - I. BPEX should be the reference
  - II. Vet should be the reference

2. <u>PMWS STUDY</u>

1. Feeling about the results

- A. Positive
  - i. Avoid relying only on vaccine
  - ii. Enhance importance of good practice (vaccines hide poor management)
  - iii. Enhance importance of good bio-security
  - iv. Bring dialog
  - v. Re-enforce past actions/decisions
  - vi. Importance to know the causes
  - vii. Useful in case the vaccines lose their efficacy (Fear that the disease strikes again)
- B. Negative
  - i. Not useful due to vaccine efficacy (not a priority)
  - ii. More interested in when to stop using the vaccine
  - iii. No new outcome, already known
  - iv. More interested in the economics
  - v. Worried about confounding

Fig. 2 Final template (continued)

Deciding the need to control

Motivations to control: Results from the quantitative questionnaire indicated that financial reasons were the most important drivers for disease control, followed by animal welfare. Image and sense of pride of being a good manager were less important, but also obtained high scores.

In the qualitative analysis, the most frequent drivers for disease control were related to the observation of ill pigs, reduced production performance (including fertility issues) and/or mortality levels. The latter was perceived as the most important factor. For most farmers, these observations were linked to economic drivers and the prospect of entering in an economically unsustainable situation. Animal welfare was also an important driver, which in combination with the economic impact, lead some farmers to a feeling of despair. This was most frequently observed in farmers describing a PMWS/PDNS disease problem experience.

Fear of disease spread within the farm or the fact that effective control measures were available was sometimes reported as enough reasons for farmers to decide to control a disease problem. In addition, reputation of the farm was an important driver for farmers selling pigs to other producers. ‘Mutual social responsibilities’ were only mentioned in the case of foodborne diseases such as salmonellosis. Other drivers for disease control mentioned by producers were explained by the impact of external influences.

External influences: In situations where external influences played an important role in the decision on whether there was a need to control or not, the veterinarian was identified as the major influence. Veterinarians frequently used data from abattoir health scores (BPHS) and farm records, and linked these data to economic performance and cost-effectiveness of control measures to support their arguments for the need to control identified disease problems. In some cases, farmers were persuaded by their veterinarians to conduct a trial (e.g. vaccination trial) to make them realize the need to control a disease problem. In other situations, pressure from abattoirs and contractors was the main driver for deciding to control a disease. Fear of diseases spreading from other farms and other producers’ negative experiences also influenced farmers to undertake preventive actions. Finally, deciding the need to control was driven in some cases by government pressure when related to foodborne diseases (e.g. salmonellosis).

Deciding how to control the disease problem

Attitudes toward measure – outcome belief: Three important nodes were identified:

Belief effective. This node relates to the belief of the farmer on the effectiveness of the control measure. Several themes composed this node and helped to explain this belief. An important theme was the ‘common sense of the control measure’, which in most cases was derived from explanation from the veterinarian. Control strategies based on good husbandry practices or drastic measures inspired farmers’ belief on the effectiveness of the control measure. In some cases, other farmers’ positive experiences were enough to convince the producer of the effectiveness of the control measure. There was also the perception that farmers from abroad and practices implemented abroad were better:

> “Again, I have been talking to other people. Contacting friends abroad. They do it on the continent all the time. Why they don’t do it over here, I don’t know. The vet has never suggested it at all. Which is common practice on the continent to do this.

> So hands…well it worked for them…and their figures are a lot better than our figures, so I will give it a go (..). I am pretty confident that is going to sort it out. Fingers crossed, yes.” (Farmer 6)

Farmer’s previous experience or the confidence and trust in the veterinarian’s guidance were also identified as factors related with farmers’ belief of the effectiveness of control measures.

Uncertainty. In many situations farmers were uncertain on the efficacy of the control measure. This uncertainty was in some cases due to the novelty of the disease affecting the farm, and the fact that no successful control measures had yet been described. In some situations, not being able to identify the cause of the disease problem or the fact of dealing with complex multi-factorial disease situations also led to uncertainty.

Principles. A few farmers reported some principles that need to be considered when deciding which control measure to implement for disease control purposes. One participant stated that, as a ‘philosophy’, antibiotics were only used when no other measures were available. Some other farmers reported, based on animal welfare issues and amount of extra labour required, that they tried to avoid injections as much as possible. The belief on the cost-effectiveness of a control measure was also reported as one of the requirements essential to some of the participants.

Subjective norms:

Subjective beliefs. Veterinarian’s opinion and advice was generally considered as the most valuable. However, many farmers reported to use other farmers’ experiences to decide on how to control ill-defined or ill-structured disease problems.

> “Usually word of mouth. It was usually at meetings and you talk to them and they said ‘well, yeah, it made a big difference to our bank balance after two years’. That persuasion, really, rather than the vet. The vet was sort of ‘well we have enough fund, do it while [making] the building that we had to do’, and it was that sort of….The vet put the seed of an idea there, but it was talking to other farmers that actually convinced us of what to do.” (Farmer 16)

Many times, other farmers’ experiences were channelled through the pig press and/or BPEX through pig discussion groups and workshops. Breeding companies, feed representatives and nutritionists were also considered as important external influences.

*Motivations to/not to comply***.** One of the most important drivers identified was related with the ‘economics’ of the control measures. Most farmers considered the cost and cost-effectiveness of the control measure as essential criteria for decision making. As shown in previous comments, relating to other producers’ positive experience with the control measures and/or explaining and making farmers understand the logic of the measures and/or the causes of the disease, were also important drivers to comply with advice. Feasibility of implementation of the control measure and amount of extra labour required for this purpose were also important drivers. The idea of positive externalities, such as additional elimination of other diseases, was also reported. In some cases, farmers finding themselves in difficult situations (‘situation of despair’) could act against their veterinary advice (social norm) and follow other farmers’ recommendations:

> "Once I knew there was a vaccine for APP…we have always been told by our vet that the vaccine for APP was not a very successful solution. So, therefore, we knew there was a vaccine available, but we understood it wasn't very efficient, so consequently it hasn't been done. He has never recommended it and he did not recommended it this time, but the situation is that we are just again back against the wall with very ill pigs, and a tremendous amount of work to try to stop them going backwards and dying. And when I heard that from my Irish friend, then I persuaded more and the second my vet came back of holiday (…) I bought the stuff." (Farmer 8)

Nevertheless, the situation itself was at times enough motivation for farmers to follow veterinary advice, without having evidence or being uncertain of the level of effectiveness. In some situations, the lack of effective measures was also a motivation to agree with external advice. However, in many situations trust in the veterinarian was an important factor for compliance by farmers with the advice received.

Perceived behavioural control:

Power to control. Three major themes were identified concerning factors limiting or facilitating the decision making process. The first one was associated with the economical situation of the farm. Farmers with an economically difficult situation or with tight budgets reported not to be able to employ the labour force needed to effectively implement control measures and/or did not have the resources to make major investments. The second major theme was related to the structure, management and environment of the farm. Several farmers believed that having strong bio-security measures or having good husbandry practices in place facilitated the effectiveness of the control measure. In addition, the design of the farm buildings was also seen as a limitation or facilitator for the implementation of different control measures. The third theme was related to the lack of existent knowledge or understanding of the disease. This was particularly felt with novel diseases, where not enough information and/or effective control measures were available. In other cases, the fact of not being able to determine the cause or understand complex multi-factorial disease situations were also identified as limiting factors.

Factors influencing the use of information sources

General issues with information: Several third order themes were identified:

Lack of effective communication. This theme was related to the excessive confidentiality existing within the pig industry that acted as barriers to knowledge. Many farmers believed that most producers are too 'individualist', and that they do not tend to share the disease problems at their farms. They believe that most farmers 'only tell the good things and not the bad things'. Not sharing problems and experiences was seen as an important issue. One participant defined this behaviour as a 'mindset of British farmers', indicating a cultural explanation. Excessive confidentiality was also an issue observed with breeding companies and the research community, which was seen sometimes as a 'danger' for disease spread. Overall, this lack of communication within the industry was described by one farmer as a feeling of isolation, which could be harmful in disease situations. Farmers' discussion groups and pig health schemes/clubs were seen as the most preferred methods of communication of disease issues by participants.

Poor communication format. Several farmers expressed difficulties to read and understand some of the information available. These difficulties were due to the excessive scientific terminology/ approach of some of the material, or to the difficulties to access information of some websites.

Lack of information. This theme related to the information missing or not communicated to farmers. One participant expressed that solutions were not generally provided/proposed with information provided on health status/diagnosis. Also mentioned were the importance of information from foreign producers, as pigs, and therefore diseases, are traded internationally, and the perception that limited research is currently being performed in the UK and hence the need to access international information.

Information bias. Some farmers raised concerns on the objectivity of the information received. Conflicts of interest from pharmaceutical companies were also noted. One participant explained that some of the information provided was derived from other farmers' experience with different systems, and therefore not applicable to his farm. This farmer also claimed that some wrong advice was supported by the media or some organisations due to the eloquence of some farmers.

Attitudes and perceptions towards information sources: Results from the quantitative survey showed that Pigworld magazine, BPEX and BPHS were identified as the most useful disease information sources by participating farmers.

Research at universities. Although most of the farmers stated that research is 'probably good' or 'that it is good that research is being done', only few farmers were aware of specific research conducted and several negative themes were obtained. The major criticism was associated with the lack of communication of research findings. For some farmers, the lack of communication or research feedback made them unable to provide an opinion about its usefulness and/or made them question the relevance of the research. One farmer criticized the fact that most of the research is not in open access journals and therefore not accessible. Another participant claimed the right of producers to have access to research findings, particularly of projects funded by BPEX. A few producers attributed this lack of awareness of research to the lack of time, of motivation, of interest, or to the fact that they do not know where and how to obtain this information. The lack of pig research funding in the United Kingdom was also perceived by farmers as one of the reasons why they did not hear about UK based research projects. In their opinion the English pig industry mainly relies on research done abroad, which in some cases is not applicable to the UK pig production. One participant reported that 'any real research' was only done abroad.

Another problem identified by participants associated with information on research projects was the poor communication format in which scientific articles normally are written. Information bias was one of the worries that some farmers had regarding university research. Three producers believed that the conditions of the experiments/trials performed by academics did not reflect the real pig farm conditions, and therefore the results were not applicable or useful to them. Two farmers specifically criticized the low sample size of some of the trials conducted. Another two participants believed research to be mainly funded by pharmaceutical companies or by other commercial companies. This implied a bias towards the interest of these companies. Some producers perceived the research to be outdated in nature, with results only available after farmers had already found a solution through other means. Another farmer believed researchers lacked field expertise and this made him reluctant to participate in research projects.

Farmers' attitudes/perceptions towards the results of the PMWS risk factor study

Positive: Many farmers believed the results to be useful as these highlighted the importance of good practice/management and farm bio-security, which in some cases can be masked by the effect of the vaccine. It was the fear of loss of vaccine efficacy or to avoid relying on a 'tablet or a vaccine' that made these results useful for some of the participants. Also it was considered useful to know that management and good farming practices can be effective alternative control measures to vaccination. The importance of understanding the cause of the disease was also identified as one of the reasons why some producers found these results useful. However, some of these comments were also related to the fear that the disease could strike again on their farms. Some producers found the results interesting as it agreed with some of their own experiences with the disease. For some farmers these results re-enforced the decisions and actions taken in the past. Farmers also reported they would keep selecting these disease control strategies if required in the near future.

Negative: Although most of the farmers believed that these results were interesting, the fact that effective PCV2 vaccines are available made these findings not as relevant or useful as they could be. In addition, some farmers stated that these results did not bring up any new information, that these were things they already knew, had found out by themselves, or related to code of good farming practices. Some of the producers reported to be more interested in knowing when they would be able to stop using the vaccination or on the economic value of the research findings. Finally, one producer became sceptical about the results due to the possible confounding involved (e.g. outdoor/indoor effect).

DISCUSSION

This study has aimed at improving the understanding of the factors involved in disease control decision-making process, information sources and management of information by farmers. Template analysis proved useful to capture the high variation of experiences and perceptions amongst farmers, but also to identify common and shared themes. It was relevant for this study to focus only on diseases classified as ill-structured or ill-defined. Application of scenarios based on diseases easy to recognize and/or to control would have failed to capture the full process of the decision-making and could have inevitably resulted in obvious, non-rich and non-diverse answers by participating farmers.

In the case of deciding how to control a disease problem, the theory of planned behaviour was used as a framework. This model presents several limitations and has been criticized by several authors. Some authors argue that in particular cases attitude may not be directly related to behaviour, while other authors criticized that different components might have different weights (Sheppard et al., 1988). Nevertheless, several empirical studies have shown the usefulness of this theory to predict and explain behaviour and it has been widely used in the agricultural sector (Ajzen, 1991; Armitage & Conner, 2001; Garforth et al., 2004; Jackson et al., 2006; Gunn et al., 2008). For this study, the framework was used to classify and organize the factors into meaningful categories that could be used to facilitate the interpretation of the findings.

‘Economics’ was one of the major themes identified consistently throughout the template. The current difficult economic environment limits and even nullifies in some circumstances farmers’ profits and their capacity to undertake important investments and implement control measures, acting as a barrier to behaviour intention linked to disease control. It provides producers with a feeling of pessimism and the constant need to consider the cost of disease and cost-effectiveness of control measures throughout the whole decision-making process. Therefore, veterinarians and other actors have to tackle and consider economic aspects, impact of disease and control measures in order to provide advice. However, economic aspects were not the only relevant factor involved in the decision process. Factors such as understanding the cause of the disease, the logic and common sense of the control measure, farmers’ principles or fears, feasibility of the control measure and animal welfare, amongst others, need also to be considered. Different farm types and sizes as well as a farmer’s personality traits might have an influence in the relevance of drivers and goals involved in the decision-making process (Willock et al., 1999).

In general, lack of communication within the industry, and especially between farmers, was seen as one of the most important barriers to knowledge. This lack of communication provides some farmers with a feeling of isolation. This is in line with findings of a survey conducted by BPEX in 2007 and with other social studies (Heffernan et al., 2008; Ellis-Iversen et al., 2010). Nevertheless, producers stated that farmers discussion groups and sharing of information between them, are crucial and the most preferred methods of communication in order to tackle ill-defined and/or ill-structured diseases.

Farmers reported an important lack of awareness and communication of research findings from academia and other institutions. It is possible that findings from academic research are communicated by the veterinarians or BPEX to farmers, but it may not to have been associated with the research bodies as such. However, although most producers considered research to be important, farmers stated that they did not seek for knowledge on research. Most of the participants expected their veterinarians, BPEX or the press to inform them about it. In addition, several negative attitudes toward research were also identified. As a result, optimal use of scientific evidence-based information may not be achieved by farmers, which could therefore compromise the effectiveness of the research done. Indeed, this study identified the lack of knowledge as an important factor in the decision-making process. Other studies have shown that lack of knowledge was one of the reasons why farmers did not implement biosecurity measures, vaccination programmes or adopt new technologies on their farms (Jonsson & Matschoss, 1998; Garforth et al., 2004; Delabbio, 2006; Heffernan et al., 2008). This lack of knowledge in some cases could be also due to attitudes towards specific information sources.

BPEX, Pigworld and, especially, veterinarians were identified as the most used and trusted sources of information. Many farmers rely exclusively on them, and sometimes just on the veterinarian, to update them on any type on of information related to disease, coming from research, area disease alerts, and any other relevant disease information sources. However, some farmers also reported the failure of these sources in updating them on disease information. Further, in some non-sustainable disease situations where no known effective control measures are available, other producers, and especially those from abroad, seemed to influence considerably farmers in their decision-making process. The results of this study highlight the importance of the veterinarians’ role and responsibilities on pig health in the farms, but also on keeping farmers updated in relation to research and disease issues.

Explanation of the PMWS severity and risk factors research findings was a useful technique to understand farmer's preference of communication of research results. In this regard, results should be communicated in combination with economic data. It should also focus on the re-enforcement of farmers past actions when dealing with this disease. The results should highlight the importance of knowing the causes of the disease and have to provide support so that farmers are aware of other effective control measures and do not rely solely on the PCV2 vaccine for disease control. It is however important to note that the researcher who conducted the PMWS research was also the one that explained the results and conducted the interviews to farmers, and therefore some of the answers obtained in this part, and on the research section, might have been influenced by a willingness to please or not to hurt the researcher feelings.

ACKNOWLEDGEMENTS

The work was funded by The Bloomsbury Consortium and by a grant (BB/FO18394/1) from the BBSRC CEDFAS initiative, BPEX Ltd, Biobest Laboratories Ltd., Pfizer Animal Health Ltd. Special thanks to all the farmers who participated in this study. We are thankful to Mr. Christopher Browne for his help in the recruitment of two farms for this study. We would like also to acknowledge Professor Dirk Werling and Professor Katharina D. C. Stärk for their contribution in setting up this project and for valuable discussions.

REFERENCES

Ajzen, I. (1991). The theory of planned behavior. Organizational Behavior and Human Decision Processes 50, 179-211

Alarcon, P., Velasova, M., Mastin, A., Nevel, A., Stark, K.D. and Wieland, B. (2011). Farm level risk factors associated with severity of PMWS. Prev. Vet. Med. 101, 182-191

Anonymous (2007). Budget: €1 935 million (2007 - 2013). European Commision. [online consulted: http://ec.europa.eu/research/fp7/index_en.cfm?pg=food]

Anonymous (2011). Defra's Evidence Investment Strategy: 2010-2013 and beyond: 2011 update. DEFRA report. [online consulted: http://ec.europa.eu/research/fp7/index_en.cfm?pg=budget]

Armitage, C.J. and Conner, M. (2001). Efficacy of the theory of planned behaviour: A meta-analytic review. Br. J. Soc. Psychol. 40, 471-499

Delabbio, J. (2006). How farm workers learn to use and practice biosecurity. Journal of Extension 44, 6FEA1

Ellis-Iversen, J., Cook, A.J., Watson, E., Nielen, M., Larkin, L., Wooldridge, M. and Hogeveen, H. (2010). Perceptions, circumstances and motivators that influence implementation of zoonotic control programs on cattle farms. Prev. Vet. Med. 93, 276-285

Garforth, C., Rehman, T., McKemey, K., Tranter, R., Cooke, R., Yates, C., Park, J. and Dorward, P. (2004). Improving the design of knowledge transfer strategies by understanding farmer attitudes and behaviour. J. Farm Manag. 12, 17-32

Gasson, R. (1973). Goals and Values of Farmers. J. Agr. Econ. 24, 521-542

Gunn, G.J., Heffernan, C., Hall, M., McLeod, A. and Hovi, M. (2008). Measuring and comparing constraints to improved biosecurity amongst GB farmers, veterinarians and the auxiliary industries. Prev. Vet. Med. 84, 310-323

Heffernan, C., Nielsen, L., Thomson, K. and Gunn, G. (2008). An exploration of the drivers to bio-security collective action among a sample of UK cattle and sheep farmers. Prev. Vet. Med. 87, 358-372

Iles, M. (2003). Livestock KT: Evaluation of effectiveness. Presentation at the LINK Sustainable Livestock Production programme Knowledge Transfer Workshop, 4 June, at Putteridge Bury Conference Center, Luton University, Natural Resources Institute

Jackson, E.L., Quaddus, M., Islam, N. and Stanton, J. (2006). Hybrid vigour of behavioural theories in the agrobusiness research domain. Is it possible? J. Int. Farm Manag. 33

Jonsson, N.N. and Matschoss, A.L. (1998). Attitudes and practices of Queensland dairy farmers to the control of the cattle tick, Boophilus microplus. Aust. Vet. J. 76, 746-751

King, N. (2004). Using template in the thematic analysis of the text. In: Essential guide to qualitative methods in organizational research. SAGE Publications Ltd., London

Noremark, M., Lindberg, A., Vagsholm, I. and Sternberg Lewerin, S. (2009). Disease awareness, information retrieval and change in biosecurity routines among pig farmers in association with the first PRRS outbreak in Sweden. Prev. Vet. Med. 90, 1-9

Sheppard, B.H., Hartwick, J. and Warshaw, P.R. (1988). The theory of reasoned action - a meta-analysis of past research with recommendations for modifications and future research. J. Cons. Res. 15, 325-343

Willock, J., Deary, I.J., McGregor, M.M., Sutherland, A., Edwards-Jones, G., Morgan, O., Dent, B., Grieve, R., Gibson, G. and Austin, E. (1999). Farmers' attitudes, objectives, behaviors, and personality traits: The Edinburgh study of decision making on farms. J. Vocat. Behav. 54, 5-36

OYSTER FARMERS' PERCEPTION OF SPAT MORTALITY OUTBREAKS:

MORE A FIRM THAN A FARM ISSUE

M. CARLIER, J. PROU, D. MILLE AND C. LUPO*

SUMMARY

Since 2008, the French oyster farming industry has experienced increased mortality outbreaks in spat of Pacific oysters. Little is known about the oyster farmers' perception of the spat mortality and their management responses. In this study, 49 oyster farmers of Charente-Maritime, which is the main spat collection sector in France, were randomly selected and interviewed in 2011. In-depth semi-directed interviews were conducted to collect data about spat mortality description, perceived causes, opinions towards control measures to implement and individual actions actually implemented. The triploid oysters and their management were mainly perceived as an indirect cause of spat mortality outbreaks. The results of this study revealed a gap between the collective intentions and the individual behaviour of oyster farmers. Although oyster farmers believed in a collective ideal towards improving farming practices, their individual improvement strategies were actually driven by economics with the aim of a short-term profit for their firm.

INTRODUCTION

In the past recurrent reductions of French oyster production were repeatedly observed due to severe mortality outbreaks associated with pathogen detections. In the 1970s, flat oyster populations, *Ostrea edulis*, almost disappeared from France while two parasitic diseases due to *Marteilia refringens* and *Bonamia ostreae* were described (Grizel, 1985). Similarly, in the late 1960s and early 1970s, the Portuguese oyster, *Crassostrea angulata*, almost disappeared from European Atlantic waters while irido-like virus infections were concomitantly described (Comps et al., 1976; Renault, 2008). Then, the introduction of a new species, the Pacific oyster, *Crassostrea gigas*, enabled the production level to recover. However, mortality outbreaks in larvae and spat were repeatedly reported in the 1990s and the Ostreid Herpes virus OsHV-1 was identified (Nicolas et al., 1992; Renault et al., 1994; Garcia et al., 2011). Since 2008, the French oyster farming industry has experienced increased mortality outbreaks in spat of Pacific oysters (Miossec et al., 2009), in association with the systematic detection of a newly described genotype of OsHV-1 (European Food Safety Agency, 2010; Segarra et al., 2010; Peeler et al., 2012).

---

* Coralie Lupo, French Research Institute for Exploitation of the Sea (Ifremer), Avenue Mus de Loup, 17390 La Tremblade, France. Email: clupo@ifremer.fr

A multi-disciplinary approach was applied to improve the understanding of the latter spat mortality process, to find ways of reducing the impact on oyster production, and ultimately, to provide information for guiding decision process. Studies were conducted mainly in the fields of biology and environment (Samain & Mc Combie, 2008; Cochennec & Baud, 2011), aiming at improving spat survival or reducing spat mortality. Resulting recommendations mainly focused on genetic selection (Degremont, 2011) and improvement of husbandry practices (Pernet et al., 2012).

However, oyster farmers' perceptions toward spat mortality and their management strategies to reduce adverse consequences have received little attention in the biological and environmental sciences or economics literature. Such understanding could help to formulate the proper policy to expand and preserve the sustainability of oyster production. The incorporation of social sciences and methodology with traditional epidemiologic approaches can provide a useful tool for exploring producers' intentions to disease control and identifying the attitudes and beliefs that influence those intentions. This or similar approaches have been used to examine animal health-related risk perception and management in pig producers (Hernandez-Jover et al., 2012) or sheep producers (Elliott et al., 2011). In the aquaculture field, only economic-related risk perception and management responses have been explored in Norwegian salmon producers (Bergfjord, 2009), Bangladesh shrimp producers (Ahsan & Roth, 2010) or Vietnamese catfish farmers (Le & Cheong, 2010). Therefore, this study was the first to investigate the oyster farmers' perception of the spat mortality, their opinion towards the control of the mortality outbreaks, and the actions they had actually implemented to deal with them, using a qualitative methodology.

MATERIALS AND METHODS

Study sample

The study population consisted of the farmers producing Pacific oysters, *Crassostrea gigas*, located in the main region of oyster production in France (*i.e.* Charente-Maritime), which contains one-third of the oyster farms (Agreste, 2005). In particular, this is the main spat collection region, supplying all other shellfish growing regions with spat (Buestel, 2009). Oyster farmers located in other regions and leasing grounds in Charente-Maritime were excluded from the study.

The study was of cross-sectional design and the sampling unit was the oyster farmer. The opinions, perceptions and attitudes towards spat mortality were assumed to be influenced by the farm activity. Thus, a random sampling strategy, stratified on the farm activity, was chosen to capture the greatest diversity in the sample. Each farmer was classified as belonging to one of the following five strata: (1) producer farmers, providing other farmers with oysters at different rearing stages; (2) market farmers, selling all of their production for human consumption; (3) sender farmers whose own production represented 70% of the sold quantity; (4) senders whose own production represented less than 30% of the sold quantity; (5) diversified farmers, rearing oysters and other shellfish species. In each strata, farmers were randomly selected (by lottery) among the list of licensed oyster farmers for leasing grounds, obtained from the Departmental Direction for Territories and Sea.

In the absence of a clear link between the number of interviews required and the objective of diversity in a qualitative survey, the sample size calculation was based on logistic and economic constraints (time and resources), to obtain a total sample size of about 60 oyster farmers, equally distributed among the five sampling strata.

Data collection

Oyster farmers were contacted by mail then by phone to make an appointment for data collection. This consisted on a semi-structured, face-to-face interview of about 2 hours. The interviews were carried out between April and July 2011 by the same, experienced and previously trained investigator (*i.e.* the first author). Qualitative data were collected about spat mortality description, perceived causes of spat mortality, actions which should be implemented to control the mortality and actions actually implemented by the farmer (Table 1), guided by a standardised questionnaire. All the questions were open-ended to generate descriptively rich, high quality narratives. After the initial response to each of the questions, farmers were then asked to explain their opinions. All the interviews were recorded with the oyster farmer's authorisation.

Table 1. Summary of questionnaire content, for the analysis of oyster farmers' perception of spat mortality outbreaks

GENERAL ITEMS RELATED TO THE FARM
Quantity of oysters produced per year
Quantity of oysters sold per year
Location of the rearing cycle
Farm activity
PERCEPTION OF SPAT MORTALITY OUTBREAKS
Description of spat mortality outbreaks that occurred on the farm since 2007
What is the main cause of mortality outbreaks, in your opinion?
ACTIONS WHICH SHOULD BE IMPLEMENTED TO CONTROL THE MORTALITY
What actions, in your opinion, could be done to control the spat mortality?
What actions, in your opinion, should not be done anymore?
Do you think actions should be collective?
ACTIONS ACTUALLY IMPLEMENTED ON THE FARM TO CONTROL THE MORTALITY
Do you have constraints to your farming activity?
Did you change your practices to control the mortality outbreaks?
How did you change your practices?

Qualitative data analysis

Qualitative data were analysed using contents analysis to identify thematic categories (Franzosi, 2004). All the interviews were transcribed on a dedicated thematic grid and anonymised. Interview transcripts were analysed collectively by a two-person research team comprising an ethnologist and a veterinary epidemiologist. Team members read through the transcripts initially in order to familiarise themselves with the raw data and to identify key themes and issues. An interpretative coding of the responses was used, being driven by the data itself and not by pre-determined categories (Franzosi, 2004). Responses were then grouped together by thematic categories. Saturation was reached with no new issues raised and previously-raised issues being repeated. Any discrepancies in the coding were resolved by discussion among the research team.

Thematic categories were then described in terms of frequency distribution. The χ^2 test or the Fisher's test (when the expected number of observations was less than 5), with $\alpha=0.05$, were applied to determine significant differences between the frequency distributions of thematic categories between sampling strata. Descriptive data analysis was performed using the SAS software, version 9.2. (SAS, 2010).

RESULTS

Sample

Of the 58 oyster farmers contacted, 49 agreed to take part, representing a participation rate of 84.5%, and 9 refused a visit, mainly claiming lack of time. The farm activity distribution did not show any significant variation in regard to farmer participation status (Chi² test, P=0.32). All participants agreed to have the interview recorded. The characteristics of the oyster farmers interviewed are summarised in Table 2. All the five farm activities were roughly equally represented. The farmers interviewed were 48 years old on average, with a range from 30 to 68 years. The majority were men. Half of the rearing cycles were located only in Charente-Maritime. Most farms included in the sample were medium sized.

Oyster farmers' perception of the mortality process

Nearly all the interviewed farmers (93%) were concerned by the spat mortality outbreaks.

Farmers' description of a spat mortality outbreak: Most of the oyster farmers described a mortality outbreak by qualitative observations such as nauseating odour ("*it stinks*") or white colour of the spat collector, corresponding to the inside colour of the bottom shell remaining attached. Quantity of collected spat is often unknown and counting the dead oysters to estimate a mortality percentage was quite unusual before 2008. Oyster farmers also mentioned the mortality was more often estimated when the oysters are bought to a hatchery or another producer farmer than when they are collected in the sea. However, although farmers did not quantify a mortality outbreak using a mortality percentage on a routine basis, they perceived a high increase of the outbreaks since 2008 and a higher impact on their flocks.

"I can't tell you, maybe we lost 30 or 40%... I just recall what I've heard, but I never counted actually"

Table 2. Characteristics of 49 oyster farmers interviewed regarding spat mortality outbreaks in Charente-Maritime, 2011

Attribute	Levels	No. of farmers	Frequency (%)
Farm activity	Producer farmers	8	16
	Market farmers	11	22
	Sender farmers	12	26
	Senders	9	18
	Diversified farmers	9	18
Geographic location of the rearing cycle	Local (only in Charente-Maritime)	24	49
	Beyond (Charente-Maritime and other regions)	25	51
Gender	Male	44	90
	Female	5	10
Age	< 35 years	6	12
	35-45 years	12	25
	45-55 years	20	41
	> 55 years	11	22
Size of farm by sold quantity (tons)	Small (< 30)	14	29
	Medium (30-250)	21	42
	Large (> 250)	14	29

Perceived causes of the spat mortality outbreaks: The two thirds (30/49) of sampled farmers were aware of the existence of a pathogen ("*the virus*") in the environment, involved in the spat mortality process (*i.e.* the murderer). Figure 1 shows the factors perceived as responsible for the mortality outbreaks (*i.e.* the motive). Most farmers (quoted by 37 of the 49 sampled farmers; 76%) believed a weakening of oysters was the main factor responsible for the mortality outbreaks. The farming of triploid oysters (15/49; 31%) and a damaged environment (24/49; 49%) due to short-term anthropic pollutions or long-term global warming were identified to contribute to this weakening.

"The hatchery diploid oysters reproduce and can weaken the natural livestock"

A majority of farmers (34/49; 69%) believed a new pathogen causing mortality had appeared due to faulty husbandry practices, such as lack of control in the triploid oyster production process (22/49; 45%) or was introduced through uncontrolled movements of live oysters (5/49; 10%).

"They played God"

"Since they produce triploids problems arose. With genetic manipulations… If we had kept producing our natural spat, we would not be in this mess today"

One third of farmers (17/49; 35%) believed the intensive husbandry practices had favoured the pathogen expression. Identified practices were overstocking of oysters in the sea (8/49; 16%) and overuse of triploid oysters (10/49; 20%).

"We must not make the oysters grow too fast […]. It puts lot of energy in shell development and I think this weaken it, and eases its contamination by the herpesvirus"

"Animals concentration in the environment is an aggravating factor, this is like an influenza epidemic in Paris!"

A non-negligible part of the farmers believed the mortality was the expression of the natural extinction of the *Crassostrea gigas* species (13/49; 27%).

"We were warned that after 40 years, the Gigas *genetic strain would weaken"*

On average, interviewed farmers quoted 4 factors responsible for the spat mortality outbreaks. Farm activity was not significantly associated with these perceptions (P=0.93).

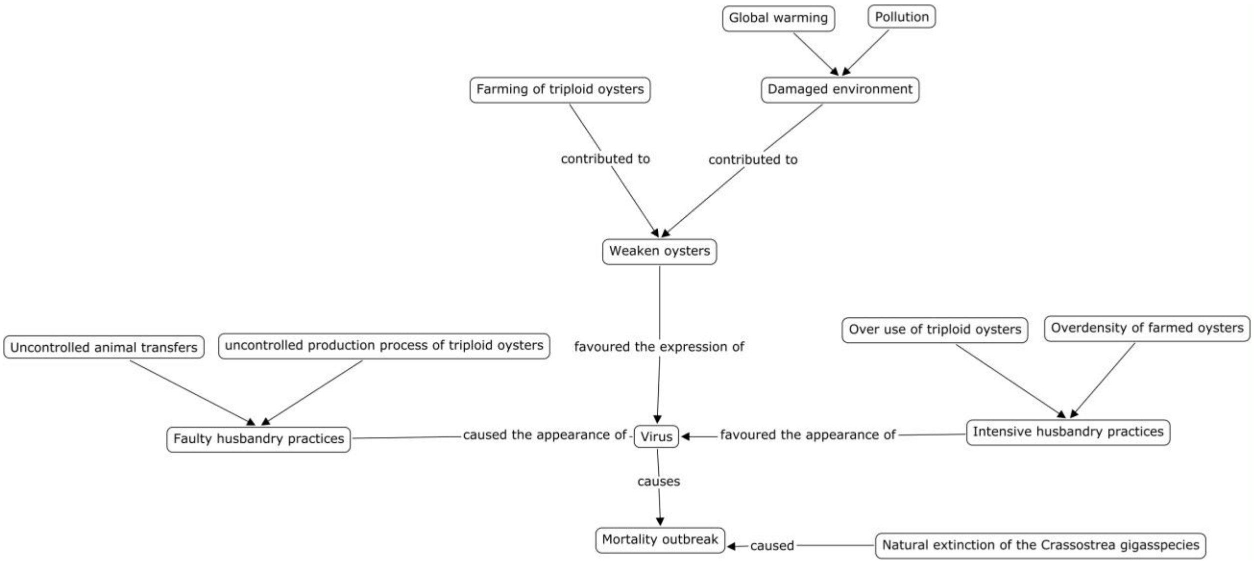


Fig. 1 Main factors perceived as responsible for the spat mortality outbreaks

<u>Actions which should be implemented to control the mortality</u>

Table 3 shows farmers' perceptions about actions likely to control the mortality outbreaks. A proportion (23/49; 47%) of oyster farmers believed there were control measures under their own responsibility. Nearly half of the farmers (23/49) believed that reducing the farming densities would contribute to contain the disease and thus, to limit the mortality.

"Intensive farming is the source of a lot of problems, I think. We should work a bit less but a bit better"

Almost one-third of farmers (14/49) believed that improvement of knowledge about husbandry practices would help to avoid weakening the oysters. A similar proportion of farmers thought about limiting the exposure of oysters to the increase of seawater temperature to avoid this weakening.

“If they [oysters] are given time to grow, to harden, to acclimatize, in the end they are trained to become more resistant”

Farmers also believed in other actions depending on other stakeholders (26/49; 53%), *i.e.* policy-makers or scientists. Farm activity was not significantly associated with these perceptions (P=0.70).

“It’s the scientists’ job to do what it must be done to fix what had been done”

One-fifth of sampled farmers (10/49) optimistically believed the best was to let nature take its course, leading to new balance. On the contrary, a quarter of farmers (12/49) adopted a fatalistic attitude without bringing into question the actual farming practices, considered as high-performing and irreproachable, nor proposing any possible control measure.

“Stop saying that we stress them!”

Of the interviewed oyster farmers, 76% (37/49) believed that collective actions were necessary but only 24% (12/49) thought they would be possible. Farm activity was not significantly associated with these perceptions (P=0.26 and 0.84 respectively).

On-farm actions actually implemented to limit or avoid the spat mortality

Three-quarters (27/35) of sampled farmers perceived difficulties to their farming activity. Most farmers (88%) had intended an action to avoid mortality outbreaks or limit the impact within their farm (Table 4). Three-quarters (37/49) of the farmers aimed at replacing the dead oysters, without any action on the perceived causes of mortality, *e.g.* by buying oysters, increasing the number of spat collectors in the sea or collecting oysters in wild beds.

“I set more spat collectors, and I try to obtain hatchery spat to offset losses”

“We do what everyone does, we set more spat collectors to try to obtain the same tonnage”

“I put more oysters in a bag in order to proportionally get more survivors”

Forty percent (20/49) changed the economic strategy of their firm, *e.g.* diversification of the products, reduction of investments or employment.

“Now we don’t keep our spat anymore: we sell it! […] we’re taking money before they die!”

Only one-third (14/49) of them changed their husbandry practices either to limit the spread of the mortality outbreaks, *e.g.* reducing the farming density, or to reduce the mortality impact.

Farm activity was not significantly associated with these perceptions (P=0.40).

Table 3. Oyster farmers' perceptions of the actions likely to control spat mortality, 49 oyster farmers, Charente-Maritime, 2011

Purpose of actions	Intentions	Practices and activities	Frequency (%)
Actions depending on oyster farming			
Containing the disease spread	Reducing the farming densities	Reducing the number of spat collectors Reducing the intensive farming	47
	Regulating the use of triploid oysters	Use of quotas Zoning the farming environment	22
	Limiting the disease spread	Retiring dead oysters from the bags Isolating infected zones Oyster movements restrictions Rearing spat in confined ponds	16
Avoiding weakening the oysters	Improving knowledge on husbandry practices	Do not manipulate oysters before the end of the mortality outbreak Do not manipulate oyster during hot period	29
	Limiting oysters exposition to the raising seawater temperature	Retiring oysters from seawater when the temperature increases Rearing oysters in deeper water	29
	Protecting the environment	Stop collecting oysters on natural beds Cleaning of leasing grounds	20
	Decreasing oyster growth	Lengthening the rearing cycle Acclimating the oysters Use of farming zones located in shallower water to have more time uncovered by seawater	16
	Increasing oyster welfare	Reducing oyster manipulations Avoiding sun-exposition of oysters on the boat Reducing mechanization for sorting oysters	16
Actions depending on policy-makers			
Regulating activities	Controlling hatcheries	Compulsory sanitary period Control of input/output of hatcheries	14
	Controlling use of genetics	Stop "*playing God*" Stop genetic experiments	12
	Protecting wild oyster beds	Regulating on-foot fishing	10
	Reducing seawater pollution by agriculture	Stopping corn culture Reducing the use of phytosanitary products	6
Actions depending on science			
Replacing the species	Importation of a new species	Recreating natural beds of wild oysters Increasing genetic diversity and limiting consanguinity Replacing actual species	27
Improving scientific knowledge	Genetic improvement of actual species	Selecting resistant oysters to the virus disease	12

Table 4. On-farm actions actually implemented to manage spat mortality, 49 oyster farmers, Charente-Maritime, 2011

Purpose of actions	Examples of practices	Frequency (%)
Offsetting spat losses	Buy oysters at different stages Collecting oysters on wild beds Increasing the number of spat collectors Increasing the farming density Increasing the farming space	76
Avoiding or limiting the economic consequences of spat mortality	Selling oysters before they die Diversification of products (other species of shellfish) Reducing investments Reducing employment Closing sites of the firm	41
Reducing spat mortality or avoiding its impact	Movement restriction Rearing spat in deeper water, or in ponds Increasing the rearing cycle to slow down the growing Reducing farming density Reducing the number of spat collectors Decreasing the number of manipulations, especially in summer Reduce mechanisation Changing farming material (e.g. spat collectors) Waiting after the mortality period to collect spat	29

DISCUSSION

This study has provided first insights on the oyster farmers' perception towards spat mortality outbreaks.

Random sampling, stratified by farm activity, ensured a representative subsample in each farm activity strata, capturing the greatest diversity of perceptions in the whole study sample. A satisfactory survey participation rate of 84.5% was achieved in this study and non-respondents did not systematically differed from respondents, limiting selection bias. Therefore, it can reasonably be assumed that the conclusions drawn from the interviewed oyster farmers reflect the spat mortality perceptions and management responses of the oyster farming industry in Charente-Maritime. However, although Charente-Maritime gathers one-third of the French oyster farms (Agreste, 2005), the present results could not be generalised to the whole French oyster farming industry. Indeed, Charente-Maritime is the traditional region for oyster farming and the main spat collection in France, with Arcachon basin (Agreste, 2005). The spat mortality perception and management may differ in other oyster farming regions, which do not produce their own spat.

In this study, oyster farmers believed the spat mortality was caused by a multifactorial process, involving a pathogen, *"the virus"*, and other factors favouring its appearance or expression such as oyster weakening. Triploid oysters and their management were often pointed out as an indirect cause of spat mortality outbreaks. Farmers believed the genetic manipulation has weakened the triploid oysters. The artificial production of triploid oysters was perceived as a possible way for appearance of a pathogen but also for dissemination in the sea because of the huge amount of spat provided by hatcheries. Part of these perceptions might be related to the history of French shellfish farming. In the 1850's, control of spat collection from the sea led to distinguish shellfish farming from fishing (Coste, 1859). Whatever the farmed oyster species, French oyster farmers have always obtained sufficient spat resource from the sea, collecting "natural" spat in summer. Hatcheries have enable oyster reproduction all around the year. In the 1990's, they started to provide triploid oysters to oyster farmers (Gérard, 1994). Traditionally, oysters are marketable for human consumption during winter. Triploid oysters are widely used because they offer better growing performances and can be sold on summer markets, offering wider economic opportunities for oyster farmers. However, their "artificial" character gives them a social deficit perception. The oyster itself, eaten fresh and alive, offers to the consumer the perception of a natural product, directly linked to its environment. Consequently, oyster farmer represents the "gardener of the sea" (Legué Dupont, 2004). Since 2008, due to the massive spat mortality outbreaks, oyster farmers developed an economic dependence to "artificial" spat provided by hatcheries, in opposition to their social and cultural history related to "natural" spat.

On the one hand, a majority of interviewed oyster farmers (69%) believed that the spat mortality had appeared or been introduced, in which they were not involved. Thus, they believed there was nothing to do to solve this, and actions actually implemented aimed at offsetting the dead spat. On the other hand, a smaller proportion (35%) believed that *"the virus"* was already present and *"something"* had favoured its expression, leading to mortality outbreaks. Consequently, those farmers' behaviour aimed at limiting the mortality impact, through a change in the husbandry practices.

In conclusion, the results of this study revealed a gap between the collective intentions and the individual actions actually implemented by oyster farmers. The collective intentions were linked to the perceived causes of the spat mortality, aiming at containing the disease spread or avoiding oyster weakening. Surprisingly, actions actually implemented by individual farmers to limit the mortality on their farms appeared to have little relation to the perceived causes of the mortality outbreaks. Although oyster farmers believed in a collective ideal towards improving farming practices, their individual behaviour actually pertained to economics with the aim of a short-term profit for their firm. The drivers influencing these collective perceptions and individual practices may be further explored to better understand and target the underlying believes. Then, these may be used to support behaviour changes more effectively than simply dictating behaviour changes, in order to improve the compliance of oyster farmers with recommendations.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the Comité Régional de la Conchyliculture Poitou-Charentes, the Department of Charente-Maritime and the Région Poitou-Charentes for funding this study, and the 49 oyster farmers for their participation in the survey.

REFERENCES

Agreste (2005). Recensement de la conchyliculture 2001. In: Direction des pêches maritimes et de l'aquaculture (Ed.) Service central des enquêtes et études statistiques, Paris, 89

Ahsan, D.A. and Roth, E. (2010). Farmers' perceived risks and risk management strategies in an emerging mussel aquaculture industry in Denmark. Mar. Resour. Econ. 25, 309-323

Bergfjord, O.J. (2009). Risk perception and risk management in Norwegian aquaculture. J. Risk Res. 12, 91-104

Cochennec, N. and Baud, J.P. (2011). Bilan des surmortalités des huîtres creuses *Crassostrea gigas* depuis 2008. Bulletin épidémiologique, santé animale et alimentation 42, 2-5

Comps, M., Bonami, J.R., Vago, C. and Campillo, A. (1976). Une virose de l'huître portugaise (*Crassostrea angulata* LMK). CR. Acad. Sci. Paris 282, 1991-1993

Coste, V. (1859). Reproduction des huîtres. In: Revue des inventions françaises et étrangères - Annales des progrès de l'industrie agricole et manufacturière (Ed.), Le Génie industriel. Armengaud Frères, Paris, France, 140-143

Degremont, L. (2011). Evidence of herpesvirus (OsHV-1) resistance in juvenile *Crassostrea gigas* selected for high resistance to the summer mortality phenomenon. Aquaculture 317, 94-98

Elliott, J., Sneddon, J., Lee, J.A. and Blache, D. (2011). Producers have a positive attitude towards improving lamb survival rates but may be influenced by enterprise factors and perceptions of control. Livest. Sci. 140, 103-110

European Food Safety Agency (2010). Scientific Opinion on the increased mortality events in pacific oysters, *Crassostrea gigas*. EFSA Journal 8, 1-59

Franzosi, R. (2004). Content analysis. In: Hardy, M., Bryman, A. (Eds.), Handbook of data analysis. Sage Publications, London, UK

Garcia, C., Thebault, A., Degremont, L., Arzul, I., Miossec, L., Robert, M., Chollet, B., Francois, C., Joly, J.P., Ferrand, S., Kerdudou, N. and Renault, T. (2011). Ostreid herpesvirus 1 detection and relationship with *Crassostrea gigas* spat mortality in France between 1998 and 2006. Vet. Res. 42

Gérard, A. (1994). Spécificités des mollusques, conditions de mise en oeuvre chez l'huître creuse Crassostrea gigas à l'échelle de l'écloserie et premiers résultats de terrain. In: International Conference "Bordeaux Aquaculture '94, Bordeaux, France, 107-121

Grizel, H. (1985). Etude des récentes épizooties de l'huître plate *Ostrea edulis* Linné et de leur impact sur l'ostréiculture bretonne. Université de Montpellier, France, 145

Hernandez-Jover, M., Taylor, M., Holyoake, P. and Dhand, N. (2012). Pig producers' perceptions of the Influenza Pandemic H1N1/09 outbreak and its effect on their biosecurity practices in Australia. Prev. Vet. Med. 106, 284-294

Le, T.C. and Cheong, F. (2010). Perceptions of risk and risk management in Vietnamese catfish farming: an empirical study. Aqua. Econ. Manag. 14, 282-314

Legué Dupont, P. (2004). La moisson des marins-paysans. INRA-MSH

Miossec, L., Allain, G., Arzul, I., François, C., Garcia, C. and Cameron, A. (2009). First results of an epidemiological study on oyster (*Crassostrea gigas*) mortality events in France during summer 2008. In: International Symposium on Veterinary Epidemiology and Economics XII, Durban, South Africa

Nicolas, J.L., Comps, M. and Cochennec, N. (1992). Herpes-like virus infecting Pacific oyster larvae, *Crassostrea gigas*. Bull. Eur. Assoc. Fish Pathol. 12, 11-13

Peeler, E.J., Reese, R.A., Cheslett, D.L., Geoghegan, F., Power, A. and Thrush, M.A. (2012). Investigation of mortality in Pacific oysters associated with Ostreid herpesvirus-1 mu Var in the Republic of Ireland in 2009. Prev. Vet. Med. 105, 136-143

Pernet, F., Barret, J., Le Gall, P., Corporeau, C., Degremont, L., Lagarde, F., Pepin, J.F. and Keck, N. (2012). Mass mortalities of Pacific oysters *Crassostrea gigas* reflect infectious diseases and vary with farming practices in the Mediterranean Thau lagoon, France. Aquacult. Env. Interac. 2, 215-237

Renault, T., Ledeuff, R.M., Cochennec, N. and Maffart, P. (1994). Herpesviruses associated with mortalities among Pacific Oyster, *Crassostrea-Gigas*, in France - Comparative Study. Rev. Med. Vet.-Toulouse 145, 735-742

Renault, T. (2008). Shellfish viruses. In: B.W.J. Mahy and M.H.V. Van Regenmortel (Ed.), Encyclopedia of Virology, 560-567

Samain, J.F. and Mc Combie, H. (2008). Summer mortality of Pacific oyster *Crassostrea gigas*. The Morest project. Editions Quae Versailles, France

SAS (2010). SAS OnlineDoc 9.2 SAS Institute Inc., Cary, NC, USA

Segarra, A., Pepin, J.F., Arzul, I., Morga, B., Faury, N. and Renault, T. (2010). Detection and description of a particular Ostreid herpesvirus 1 genotype associated with massive mortality outbreaks of Pacific oysters, *Crassostrea gigas*, in France in 2008. Virus Res. 153, 92-99

UNDERSTANDING KNOWLEDGE, ATTITUDE AND PRACTICES (KAP$_S$) OF CAMBODIAN SWINE PRODUCERS ON PRRS AND RELATED CONTROL OPTIONS, DURING DISEASE ABSENCE AND OUTBREAK EVENTS

B. TORNIMBENE*, V. CHHIM, S. SORN, T.D. DREW AND J. GUITIAN

SUMMARY

Cambodia experienced its first reported outbreak of porcine reproductive and respiratory syndrome (PRRS) in 2010, which was characterized by high morbidity and mortality rates. The aim of this study was to better understand Cambodian swine producers' behaviour in relation to PRRS control, and what might alter their willingness to adopt good practices and disease prevention.

A KAPs (Knowledge, Attitude and Practice) survey is a representative study of a specific population to collect information on what is known, believed and done in relation to a particular topic. A survey of the KAPs of pig owners with regards to PRRS and other respiratory diseases was conducted in semi-commercial and backyard farms in Takeo province in Cambodia. The survey was designed to record data on knowledge of PRRS disease and its transmission, farmers' attitudes and practices towards control measures, knowledge on vaccination and perception towards local veterinary authority activities.

INTRODUCTION

Porcine Reproductive and Respiratory syndrome (PRRS) was first detected in Cambodia 2010 and the disease was responsible for high morbidity and high mortality in adult pigs, costly impacting on affected farmers livelihoods.

PRRS is caused by PRRSV, a single stranded RNA virus, for which two genetically different strains are recognised: Type I, initially isolated in Europe, and Type II, initially isolated in North America (Murtaugh et al., 1995; Nelsen et al., 1999). The virus is generally responsible for reproductive failure in sows, slow growth in adults pigs, and death in young pig due to respiratory diseases often aggravated by secondary infection (Cho & Dee, 2006). In 2006 a new highly pathogenic variant of the Type II strain has emerged in China: since then it has devastated the pig population of China and Viet Nam, occurring also in Lao PDR, Thailand, and Indonesia (Hu et al., 2012).

---

* Barbara Tornimbene, Royal Veterinary College, Department of Production and Population Health, Hatfield, AL9 7TA, United Kingdom. Email: btornimbene@rvc.ac.uk

The aim of this study was to describe swine producers' behaviour and identify factors that might alter their willingness to adopt good practices towards prevention and control of PRRS and other respiratory diseases of swine. The survey was designed to record data on knowledge of PRRS disease and its transmission, farmers' attitudes and practices towards control measures, knowledge on vaccination and perception towards local veterinary authority activities.

MATERIALS AND METHODS

Study design

A survey of the KAPs of pigs owners with regards to PRRS and other respiratory diseases that shows similar symptoms to PRRS, was conducted in a number of semi-commercial (SC) and backyard (BY) farms in Takeo province between March and April 2012. Farms were selected from on a sample frame generated by our team while carrying out two PRRS epidemiological surveys in the same province in 2010 and in 2011.

Takeo province was chosen as area of study based on the following criteria: first, the relatively high pig density and the importance of the region for pig production (Chetra, 2009); second, its proximity with Viet Nam, in view of the recent HP-PRRS outbreak experienced by the southern part of the country (Feng et al., 2008; Le et al., 2011); third, the coexistence of different production system, mainly medium and small scale farms (Tornimbene, 2012).

Survey methodology

Two translator/investigators from the Royal University of Phnom Penh were trained in KAP survey methodology and briefed on the characteristic of the swine farming system in Cambodia and the epidemiology of PRRS. A standardized, structured questionnaire was used to collect information on demographics (gender, age, education, profession), farms characteristics, knowledge of PRRS and its transmission patterns, farmers attitude and practices towards control measures in four different diseases scenarios (absence of disease and three outbreaks events: in district, in village, in farm), knowledge on vaccination and perception towards village animal health workers (VAHWs) and local veterinarians' activity.

Interviews were undertaken in Khmer language using a dialogue-based format. Questions were mostly in open-ended format. The interviewer went through each question without prompting any response. Likert-scale items were used to assess the level of agreement of the interviewee with a number of knowledge queries (WHO, 2008). Ranking cards were used to evaluate farmers' attitude towards control measures in different diseases scenarios: each card reported a different control practice which the interviewee had to put in order of importance in relation to the different scenarios.

Data analysis

Data from the questionnaire were translated into English, compiled in Excel (Microsoft, Redmond, WA) and quality was controlled for entry errors. Descriptive statistics were used to summarize qualitative data and frequencies were obtained in order to assess differences in KAPs between SC and BY farms.

Regression models were used to analyse the association between seven independent variables and a number of different outcomes: knowledge of disease, knowledge of viral transmission, use of vaccine and farmers' attitude towards VAHW and public veterinarians.

Independent variables considered were: sex of the respondent, years of schooling, duties in relation to pig rearing, previous experience of PRRS outbreak, herd size and farming system (farrow to finish, farrow to finish and breeder, farrow to finish and fatteners, breeder, fattener) and awareness of PRRS signs being caused by an infectious disease. SC and BY farmers' responses were merged; farm typology (SC or BY) was considered a potential confounder and incorporated as a fixed effect in each of the models.

Due to the hierarchical structure of the data obtained (farmers - level one- nested within villages – level two - nested within communes – level three) multilevel regression was applied. Linear mixed models (GLMMs) were implemented in STATA 11.2 (StataCorp LP, TE). The adaptive approximation method was used for all models, to improve accuracy and calculation time (Rabe-Hesketh et al., 2004; Rabe-Hesketh & Skrondal, 2012).

Communes and villages were initially added to the models as random effects. Maximum likelihood (ML) estimates of parameters were obtained for models with and without random effects containing only a constant, and likelihood ratio tests were used to assess improvements in the model fit. It was decided to retain commune level in all models even if variance estimates were not significantly different from zero, in order to allow for identical data structure through the whole analysis. However, village level variance proved always to be very small, and because there were no major differences in the magnitude of coefficients between a model with and without village random effect, it was decided to exclude this level.

Univariate analysis was performed using binary logistic or ordinal regression, depending on the outcome of interest. Variables with significance level $\alpha \leq 0.20$ where checked for co-linearity by calculating variance inflation factors (VIF), and variables with VIF < 10 were retained for further analysis (Dohoo, 2010; UCLA, 2012). Once in the final multivariate model, non-significant ($P > 0.05$) variables were removed sequentially. Biologically meaningful interactions (e.g. sex and year of schooling) were manually tested and retained when $P \leq 0.05$.

Ethical approval

Permission to undertake the work reported was granted by the Cambodian National Veterinary Research Institute, which is part of the Animal Health and Production office of the Ministry of Agriculture Food Fishery and Forestry. The questionnaire was previously submitted and approved by the RVC Ethics and Welfare Committee.

RESULTS

The survey was conducted on 112 farms, comprising 34 semi-commercial farms and 78 backyard farms respectively. Overall, 42 villages within 24 communes of 7 districts, were visited. Male farmers seemed to be more often in charge on SC farms, while a large percentage of BY farmers were women. Fattening was the most common activity between interviewees.

Descriptive analysis

Knowledge of PRRS: Once presented with the list of clinical signs characteristic to HP-PRRS in Cambodia, most farmers (almost 100% for both farm types) indicated infectious diseases as one of the major causes. Around 40% of farmers indicated weather changes as the second most likely causes; diet, enteric worms and insect bites were mentioned by approximately 20% of farmers.

When asked to interviewee how they thought the disease was moving between pigs, apart from direct contact, 53% of SC farmers and 19% of BY farmers, indicated “people who come into contact with pigs” as one of the potential routes. When farmers were then asked to respond to specific questions on transmission routes typical to PRRSV (direct contact, aerosol, fomites, dung, venereal transmission), most of them seemed confident in agreeing that those were potential routes of infections.

Interviewees were then asked how they thought the disease could spread between farms located in the same village, and respectively 74% and 45% indicated “other farmers and/or people moving between those farms” as one the main routes of spreading, together with wind (38% and 31%); transmission of the disease between farms situated in different villages, was viewed from both SC and BY farmers as caused by “other animals present in the village which have access to the pig farms” (47% and 19%), wind (29% and 24%), infected pork meat (21% and 28%) and specifically by traders and veterinarians (21% and 8%).

Attitude and practices towards biosecurity: Not many differences in perceptions between SC and BY farmers were identified: vaccination and disinfection were the most frequently cited; for the last scenario (PRRS circulating in the farms) 71% and 68% also indicated “separate sick pigs from healthy” as one of the most important measure to apply. Quarantine was never mentioned, neither in during disease epidemics nor in the absence of disease. Controlling free-range pigs, both from entering or leaving the farm, was mentioned only by a small percentage of interviewees and very few farmers thought about stopping the purchase of pigs from other farms/villages during active disease scenarios.

When asked what practices farmers physically implement when they have diseased pigs in the farm, SC and BY farmers seemed to have a very similar approach towards bio-management practices and bio-containment actions (Fig. 1). Most farmers (88% and 73%) choose to treat sick animals hoping in recovering, and respectively 62% of SC farms and 50% of BY sell pigs directly to slaughterhouse of killing point owners, instead of middlemen or traders.

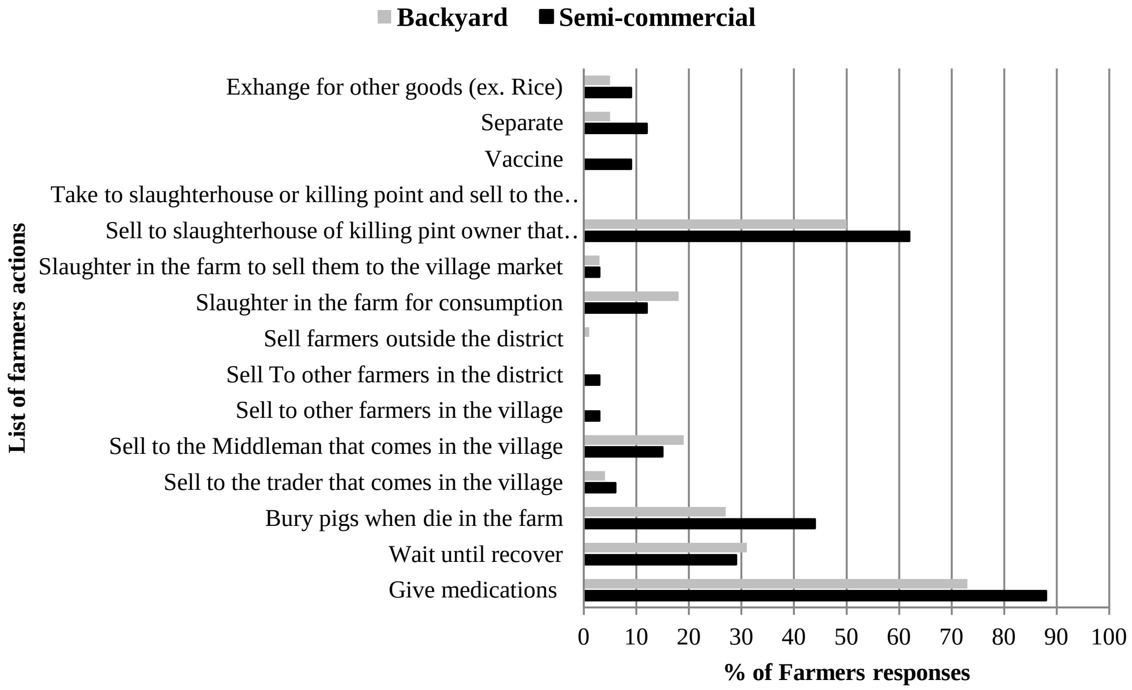


Fig. 1 Farmers' actions towards pigs showing symptomatology

Vaccine: Even if vaccination was selected as one of the most effective control measures (15% of SC and 19% of BY farmers were not using it), half of farmers were under the impression that vaccines were not always effective, but only occasionally effective; only 4% and 3% farmers raised the issue that vaccine might be difficult to use because of the absence of cold chain; also, almost no farmer thought that they were too expensive or difficult to find.

Most farmers knew that vaccines were used in order to protect pigs from diseases and some farmers (5% of BY) were also under the impression that they were helping with the growth of the animals. When asked about vaccine coverage for different diseases, 70% of farmers thought that one single vaccine type was able to protect pigs from more than one illness, or had not considered it.

Local veterinary authority: Thirty-five per cent of SC and 38% of BY farmers would call the VAHW on any occasion, and only 9% SC and 18% of BY farmers would call the local veterinarian. Respectively, 32% of SC would never call the VAHW and over 70% of all farmers would never call the local veterinarian. Overall, 70% of all farmers were confident that, if a VAHWs or a local veterinarian visited, it would have helped improving the health of all or some of sick pigs (Fig. 2).

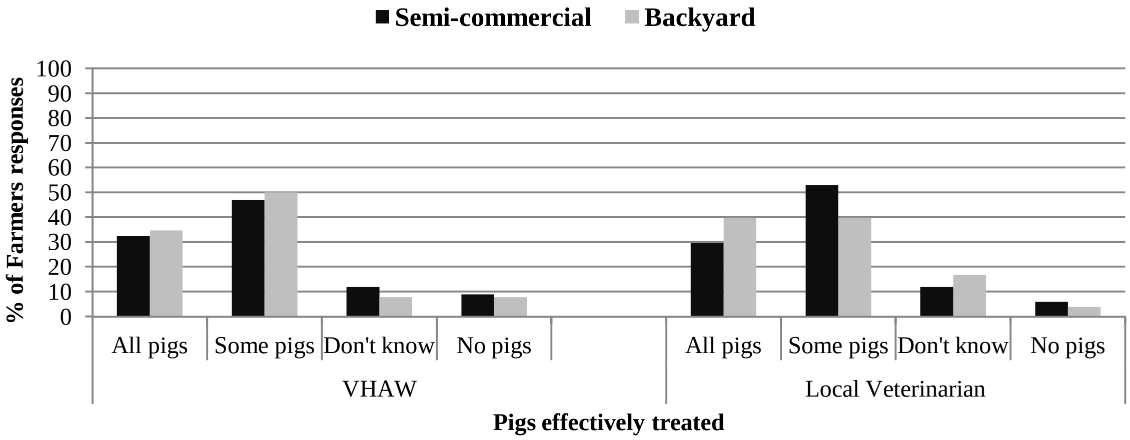


Fig. 2 Farmers' perception of VAHW and Local Veterinarian effectiveness (number of pigs effectively treated)

Of the 65% of SC and 62% of BY farmers who were not contacting the VAHW, around 30% stated that they were called out only when pigs were very sick and when medicines bought privately were not effective. Of the 81% of SC and 72% of BY farmers that where not contacting local veterinarians, more than half of farmers responded that they were not sure how to reach them or felt they were living too far from the farm.

Identification of factors affecting KAPs towards PRRS

Results from the multilevel model related to farmers knowledge of PRRSV transmission identified that male farmers and farmers with a higher level of education were more likely to have a better knowledge of transmission routes between pigs. Commune level intra-class correlation (ICC) was 0.03, signifying almost total independence between observations in within the cluster (the proportion of variance explained at commune level was only 3%).

The multilevel model associated with the knowledge of PRRSV specific transmission routes showed that male farmers and farmers with a secondary level education had higher odds of replying correctly to the aerosol and the fomites routes questions. Males seemed to be more aware of the probability of boars infecting sows, while farmers that were mostly working with crops in the field had lower odds of replying correctly to the "sow infecting boars" question. No independent variables were associated with the correctness of response about viral transmission through direct contact and swine dung. The proportion of variance explained at commune level was respectively 0% for the aerosol outcome, 43% for the fomites outcome, 24% for boar to sow transmission and 41% for sow to boar transmission.

Results of the two models analysing the association of factors to farmer beliefs on vaccination shown that BY farmers showed higher odds of knowing that one vaccine type covers for only one specific disease, and in general, farmers with this type of knowledge seemed to find vaccination more effective. No other variable was significantly associated. The proportion of variation due to commune level was 5% and 3%, respectively for knowledge of vaccine action and perception on vaccine effectiveness.

Finally, the two multilevel models of the associated with factor affecting farmers' perception towards the local veterinary authority show that the frequency call rate for

VAHWs decreases as the herd size increases. Also, the data indicates that farmers that contact VAHWs more frequently are less prone to call the local veterinarian. The proportion of variation due to commune level was 0% and 15%, respectively for VAHWs and local veterinarians.

DISCUSSION

This study successfully integrates previous knowledge of features of the disease of PRRS in Cambodia with an analysis of one of the most important component of its epidemiology - those human beliefs and actions that might shape the way the disease develops in a population. Although the results of this survey can be considered just a snapshot of a single Cambodian province, Takeo, it is true that Cambodian farming practices seems to be consistent through geographical areas with similar characteristics (Tornimbene, 2012). We are therefore quite confident that our outcomes are a good representation of farmers' behaviours in the whole lower Mekong region of the country, which is also the area with the higher pig density.

Differences between Cambodian semi-commercial and backyard farms were not too marked in terms of knowledge, attitude and practices. This reinforces our impression that semi-commercial husbandry techniques are still quite rudimentary and that SC farmers have limited understanding of disease epidemiology and application of control measures. This could prove costly to farmers in the region, considering that the country entry point for the 2010 PRRS epidemic was probably semi-commercial farms, which have larger herds, trade more widely, have a bigger turnover of pigs and raise different age groups (Dietze, 2011 and personal observations).

Although farmers did demonstrate an awareness about what might cause PRRS-like signs, this was not associated with having experienced the disease in 2010; we could infer this is a type of knowledge that might have derived from experience of endemic situation of similar respiratory diseases, which are recurrent in Cambodia (Psilos, 2008). Farmers' second most reported cause for listed clinical signs was weather changes, and although PRRS has shown an element of seasonality in Viet Nam, it is unlikely that Cambodian farmers have experienced the disease frequently enough for them to build a mental timeline. A possible explanation derives from the fact that heat raises pigs stress levels and inhibits immune responses, so that pigs tend to get sicker during the dry season (Le et al., 2011).

Besides investigating farmer awareness of the existence of infectious diseases, this project aimed to explore the level of farmers' knowledge in terms of disease biology; interviewees were left free to respond on questions about 'pig to pig' viral infection routes, direct contact and the role of people as transmitters, the latter was the sole mechanism that was frequently mentioned. However, when specifically asked about transmission routes typical to PRRS, farmers seemed quite knowledgeable; this discrepancy could be related to various factors: our open questions weren't clear or formulated correctly, farmers didn't feel the need to be very specific, farmers had difficulties in formulating precise opinions and/or think critically, or finally they didn't perceive those routes of infection important enough to be mentioned.

It is not surprising that secondary school education had a positive impact on farmers' knowledge on transmission routes; we postulate that farmers who are more accustomed to study might find it easier to grasp the concept of transmission. Interestingly, responses to questions on fomites and venereal transmission seemed to vary among communes, which may be related to different quality of schooling, or different quality of information to which farmers might have access.

In general, male farmers appear to be more knowledgeable, though an analysis for a correlation between sex and year of schooling failed to identify any significant association between the two variables. The fact that males are more likely to have a better knowledge in term of disease transmission could be related to the fact that they are mostly in charge of bigger farms spend most of their focus into pig rearing.

Interestingly, both knowledge of infectious diseases or having previously experienced PRRS did not have an impact on farmers' knowledge of transmission routes; this might demonstrate that even if farmers are aware of an event, they might only have a limited understanding of it. This raises the question of how farmers perceive the infection itself: surprisingly farmers do acknowledge other people and animals as mechanical vectors (traders, middleman and roaming animals for between villages transmission), but almost ignore the role of boars, which travel extensively from farm to farm, often covering quite long distances, and come into direct contact with sows and potentially other pigs present in the farms.

Attitude towards control measures were similar for all disease scenarios, which suggests a very limited concept of prevention as such. In general, very few farmers seemed to be concerned about biosecurity. In the event of disease, farmers tend to keep their pigs as long as possible, giving them medication and hoping for recovery, thus increasing chances of vector contamination. A limit of our study was to not differentiate farmers' actions involving pigs of different age groups, to try and better understand what are farmer practices involving young stock. We have seen that farmers treat pigs without clinical sign as healthy pigs: those might then be traded to other farms or middlemen, enhancing chances for disease spread.

Most farmers use vaccination to protect pigs from diseases; however, vaccines don't appear to be always effective. Farmers' knowledge concerning aspects that could influence vaccine efficacy, such as the specific disease target, protocols to follow and storage requirements, seemed very limited. For example, very few farmers acknowledged the lack of cold storage as a major limitation, in a country were temperatures can reach 42°C in the dry season. Surprisingly, backyard farmers were more aware than semi-commercial farmers that one vaccine type covers only for one specific disease; this might have been due to the fact that semi-commercial farmers are more likely to be using multi-pathogen vaccines, so that from their point of view one vaccine type does cover for more than one disease. There was not much variation between responses in different communes, so we can assume a small role of VAHWs or the local veterinarian in affecting farmers' behaviours in term of vaccination, thus it would be interesting to investigate the role of private drug dealers, which travel through many districts (Tornimbene, 2012).

Farmers seemed quite confident about the effectiveness of both VAHWs and local veterinarians, so we might assume farmers tend to not contact them during an outbreak for different reasons. For example, the fact that farmers have easy access to drugs and vaccines, and can afford them, might explain why semi-commercial farmers tend to rely less on VAHWs or local veterinarian and prefer treating pigs themselves; in fact, they would have to cover double costs (drugs and any fee the VAHWs or the vet charges) without being sure pigs will get better.

From discussions with farmers it is clear that they do not see the two professional figures as complementary, but tend to use one or another (generally the VAHWs who are much easier to reach). The model analysing the association between factors and the frequency with which farmers were requesting VAHWs to visit their farms in the event of disease did not highlight a big variation in terms of farmer responses in different communes, while there was some variability in responses with regard to the local veterinarian; we think this reflects the natural variability existing between professional entities. In general, farmers do acknowledge the positive function of local veterinarians, but are still unwilling to involve them, ether because of lack of habit or physical impediments (distance, unknown contact) rather than a general mistrust towards government employees.

ACKNOWLEDGEMENTS

Authors are grateful to Boehringer Ingelheim Ltd who sponsored this study as part of the PhD studentship of BT. However, this sponsor played no part in the study design, in the collection, analysis and interpretation of data; in the writing of the manuscript or in the decision to submit this manuscript for publication. We fully acknowledge the contribution of Dr. Yu-Mei Chang for her support in the statistical analysis of the data.

REFERENCES

Chetra, S. and Bourn, D. (2009). Review of animal production in Cambodia. In: FAO, MAFF Cambodia (Ed.), Environmental Animal Health Management Initiative for Enhanced Smallholder Production in South East Asia

Cho, J.G. and Dee, S.A. (2006). Porcine reproductive and respiratory syndrome virus. Theriogenology 66, 655-662

Dietze, K., Pinto, J., Wainwright, S., Hamilton, C. and Khomenko, S. (2011). Porcine Reproductive and Respiratory Syndrome (PRRS): virulence jump and persistence circulation in Southeast Asia. In: FAO (Ed.), Focus on, Vol. 5, 1-8

Dohoo, I., Martin, W. and Stryhn, H. (2010). Veterinary Epidemiologic Reserach. VER inc., Charlottetown - Prince Edward Island - Canada

Feng, Y., Zhao, T., Nguyen, T., Inui, K., Ma, Y., Nguyen, T.H., Nguyen, V.C., Liu, D., Bui, Q.A., To, L.T., Wang, C., Tian, K. and Gao, G.F. (2008). Porcine respiratory and reproductive syndrome virus variants, Vietnam and China, 2007. Emerg. Infect. Dis. 14, 1774-1776

Hu, S.P., Zhang, Z., Liu, Y.G., Tian, Z.J., Wu, D.L., Cai, X.H. and He, X.J. (2012). Pathogenicity and distribution of highly pathogenic Porcine Reproductive and Respiratory Syndrome Virus in pigs. Transbound. Emerg. Dis. 'Online fist' published 5th July 2012 [Available from: DOI: 10.1111/j.1865-1682.2011.01239.x]

Le, H., Poljak, Z., Deardon, R. and Dewey, C.E. (2011). Clustering of and risk factors for the porcine high fever disease in a region of Vietnam. Transbound. Emerg. Dis. 59, 49-61

Murtaugh, M.P., Elam, M.R. and Kakach, L.T. (1995). Comparison of the structural protein coding sequences of the VR-2332 and Lelystad virus strains of the PRRS virus. Arch. Virol. 140, 1451-1460

Nelsen, C.J., Murtaugh, M.P. and Faaberg, K.S. (1999). Porcine reproductive and respiratory syndrome virus comparison: divergent evolution on two continents. J. Virol. 73, 270-280

Psilos, P. (2008). Swine Marketing in Cambodia: Reducing market distortion, Improving competitiveness, and Ensuring safe trade. In: Limited, Indochina Research (Ed.)

Rabe-Hesketh, S. and Skrondal, A. (2012). Multilevel and Longitudinal Modeling Using Stata, Vol. 2, STATA Press

Rabe-Hesketh, S., Skrondal, A. and Pickles, A. (2004). GLLAMM Manual. The Berkley Electronic Press

Tornimbene, B. (2012). Characterisation of swine production systems in the Cambodian Mekong lowland region. In: Wordpress.com. http://agrocambodia.files.wordpress.com/2011/05/characterisation-of-swine-production-systems-in-the-cambodian-mekong-lowland-region1.pdf

UCLA (2012). Logistic Regression Diagnostic. In: Services, Academic Technology (Ed.), Stata Web Books, 2012

WHO (2008). Advocacy, communication and social mobilization for TB control: a guide to developing knowledge, attitude and practice surveys

METHODS

AN APPROACH TO EVALUATING THE RELIABILITY OF DIAGNOSTIC TESTS ON POOLED GROUPS OF INFECTED INDIVIDUALS

M. HALL[*], L.A. MUNRO, I.S. WALLACE, R. MCINTOSH, K. MACNEISH
AND A.G. MURRAY

SUMMARY

An experimental design and statistical analysis providing information on the reliability of pooled test procedures is described. It involves estimating the relationship between the probability of a pool testing positive (dependent variable) and the expected number of infected individuals in a pool (explanatory variable). The intercept is an estimate of the proportion of false positives (1-pooled specificity) and pooled sensitivities can be estimated for indicative prevalences of infected individuals. Simulations for a theoretical infection are used to investigate the advantages and limitations of the approach. The approach is also used to evaluate the reliability of a virus isolation and qRT-PCR test procedure detecting the pathogenic agent necessary for the development of Pancreas Disease in Atlantic salmon.

INTRODUCTION

Field evaluations of the reliability of veterinary diagnostic test procedures (hereafter referred to as tests) detecting infection, illness or disease (hereafter referred to as infection) are becoming increasingly available. Within the context of salmonid aquaculture, for example, evaluations of tests detecting infection with *Nucleospora salmonis* and infectious salmon anaemia virus have been published (e.g. Georgiadis et al., 1998; Nérette et al., 2005). These evaluations provide estimates of diagnostic sensitivity (Se) and specificity (Sp) as measures of the probability that a test will correctly identify either infected individuals as being infected or uninfected individuals as being uninfected, respectively. Such estimates are regarded as primary test performance indicators by the World Organisation for Animal Health (OIE) (World Organisation for Animal Health, 2009).

The pooling of samples from several individuals for a single test has long been advocated as a way of reducing the cost and effort of diagnostic testing (Dorfman, 1943). In a veterinary context pooling has been used in the identification of infected individuals (e.g. Kennedy, 2006) and populations (e.g. Kinde et al., 1996) and to estimate the prevalence of infected individuals (e.g. Raizman et al., 2011) and populations (e.g. McBeath et al., 2009). The OIE recognises the utility of testing pools (World Organisation for Animal Health, 2009, 2012) but, with regard to internationally listed diseases, stipulates that the results of such tests must

---

[*] Malcolm Hall, Marine Scotland Science, 375 Victoria Rd., Aberdeen, AB11 9DB, UK. Email: malcolm.hall@scotland.gsi.gov.uk.

be interpreted using estimates of diagnostic sensitivity and specificity for pools (World Organisation for Animal Health, 2012).

Experimental designs and statistical methods used to evaluate the reliability of tests on individuals in the veterinary context are well established (Enøe et al., 2000; Branscum et al., 2005). In contrast approaches using pools appear to be less developed. The primary purpose of this report is, therefore, to promote debate on the experimental designs and statistical analyses most suited to evaluating the reliability of pooled tests. This is done by describing a simple approach for evaluating the reliability of pooled tests and illustrating this using both a hypothetical and a real viral infection of Atlantic salmon.

OUTLINE OF THE APPROACH

This is a modification to an experimental design formalised by Hui and Walter (1980) and widely used to evaluate individual test reliabilities. The original design involves sampling individuals from two populations characterised by different infection prevalences and testing the individuals using two tests detecting the same infection. This original design has since been generalised to more than two populations and/or tests (Branscum et al., 2005). The modification involves randomly assigning individuals to groups comprising a small fixed number of individuals (pools) and subjecting those pools to the same diagnostic tests. Each individual is therefore tested for a specific infection as an individual and as a part of a pool, as illustrated in Fig. 1. It is advisable to select populations characterised by substantially different prevalences of infected individuals.

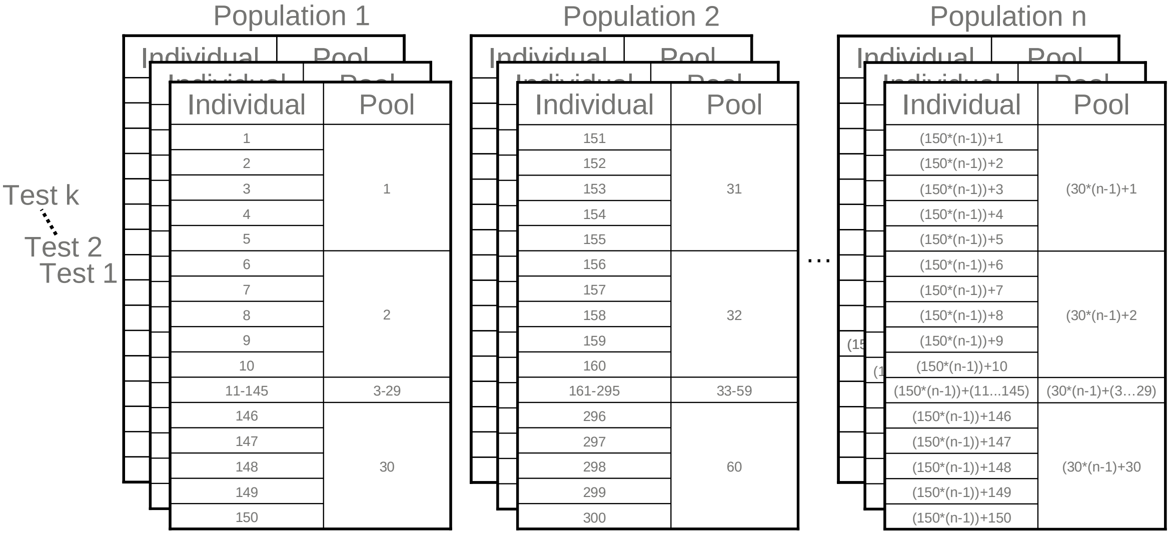


Fig. 1 Modified experimental design for estimating test reliability for individuals and pools

The statistical analysis starts with the familiar estimation of individual Se and Sp and the prevalence of infected individuals in each sampled population. Estimates of these parameters can be obtained from the posterior distributions of Markov Chain Monte Carlo (MCMC) chains generated by slice sampling (Neal, 1997) using vague priors utilising the statistical model of either Hui and Walter (1980) for tests which are assumed to be conditionally independent or, providing there are sufficient degrees of freedom, the model of Dendukuri and Joseph (2001) for tests which are likely to be conditionally dependent (Branscum et al., 2005). The subsequent analysis for pooled tests involves calculating the probability of infection for each individual using the estimates of Se and Sp and the population prevalences and enumerating the expected number of infected individuals in each pool. The relationship between the probability of a positive pooled test as the dependent variable and the expected number of infected individuals in a pool as an explanatory variable is then modelled with a Generalised Linear Model assuming a binomial error distribution and utilising the logistic link function. For this report this is fitted by penalised maximum likelihood (Firth, 1993) which, in addition to reducing asymptotic bias, is stable and provides estimates for data with complete or quasi-separation. The regression parameters are then used to predict values for PSp (1-intercept) and PSe (for non-zero integer values of the explanatory variable). The ability of a test to detect pools containing infected individuals can also be characterised by the number of infected individuals required in a pool with a PSe of 50% (ED_{50}). This process is summarised in Fig. 2.

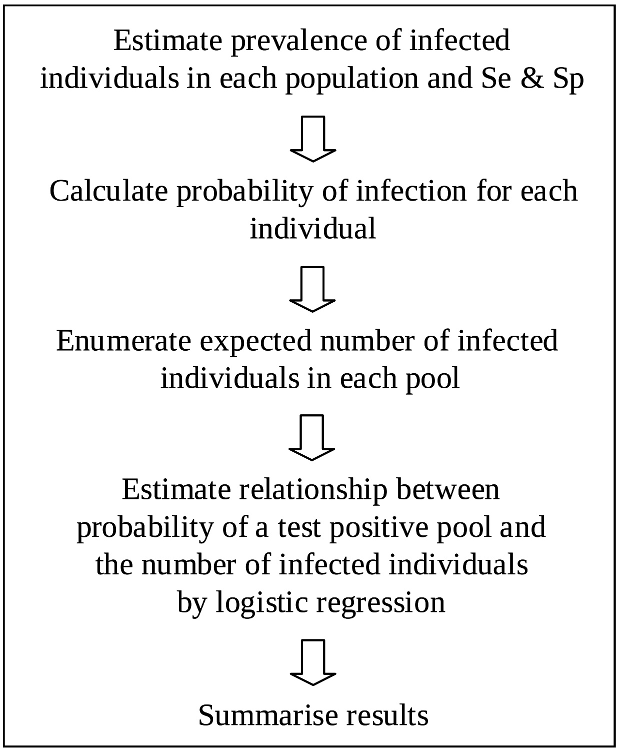


Fig. 2 Summary of statistical approach used to estimate the reliability of pooled tests

There are several problems with the approach described in the previous paragraph. One of these is that ED_{50} is not a preferred OIE test reliability parameter (World Organisation for Animal Health, 2012). This is overcome by using the OIE test validation scenarios which include testing for freedom from disease and confirmatory diagnosis of suspect clinical cases (World Organisation for Animal Health, 2009). Testing for freedom from disease requires the detection of a low prevalence of infected individuals recommended to be two percent ($PSe_{0.02}$) in the absence of reliable information (World Organisation for Animal Health, 2012). Confirmatory diagnosis of suspected infection usually involves targeted sampling of suspect individuals within a suspect population and an indicative prevalence of 80 percent ($PSe_{0.80}$) has been set by the authors. Given these indicative prevalences it is then straightforward to estimate the proportion of pools containing different numbers of infected individuals and generate estimates of PSe relevant to the problem at hand. Other problems with the approach will become apparent throughout the report and are described in the discussion.

RESULTS FOR A THEORETICAL INFECTION

Simulated data for imaginary populations containing different proportions of individuals infected with an imaginary pathogen and tested individually and in pools with imaginary tests of known sensitivity and specificity were used to validate the approach. Populations were generated by randomly assigning an infection status to a total of 900 individuals sampled in equal numbers from six independent populations characterised by individual infection prevalences of 0.02, 0.20, 0.40, 0.60, 0.80 and 0.98. Test results for individuals, conditional on infection status, were randomly generated for two conditionally independent diagnostic tests, tst_{subopt} and tst_{opt}, with Se of 50 and 85 percent respectively and Sp of 95 and $\geq$99 percent respectively. Test results for pools, conditional on the number of infected individuals randomly assigned to pools, were randomly generated for the pooled versions of the same diagnostic tests assuming PSp = Sp, PSe_{subopt} of 10, 19, 34, 52 and 70 percent for pools containing between one and five infected individuals, and PSe_{opt} of 85 and >99 percent for pools containing one and greater than one infected individual. The reliability of the pooled tests was then evaluated as described in the previous section. For the purposes of this report a suboptimal test was characterised by a Sp and PSp of less than 99 percent, a Se of less than 80 percent and a $PSe_{0.02}$ of less than 80 percent. Results for a 'typical' simulated dataset are presented in Fig. 3 and Table 1.

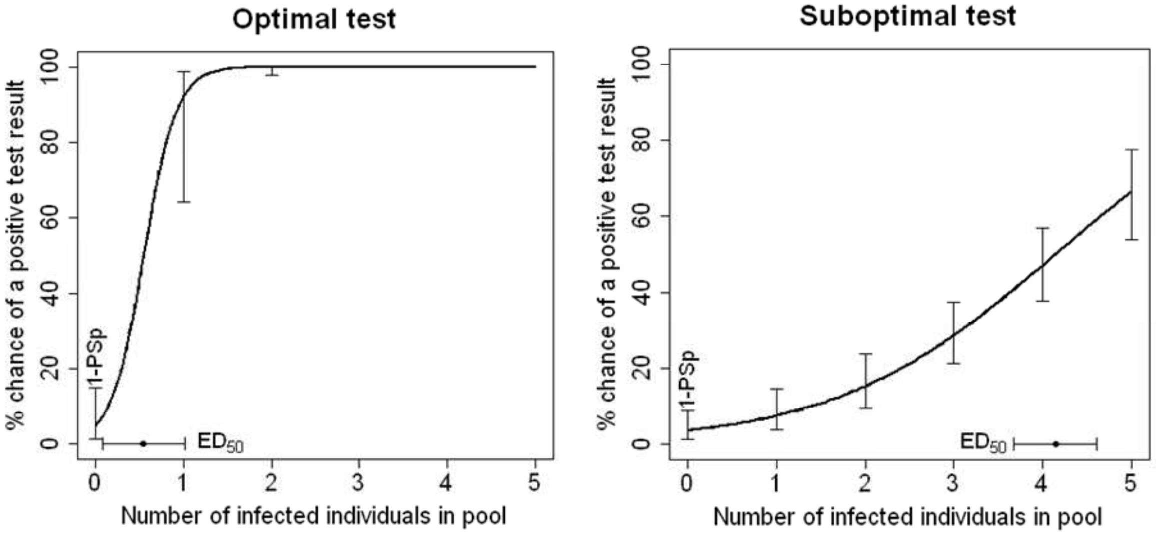


Fig. 3 Plots of predicted reliability for pooled tests

Table 1. Summary of predicted reliability parameters for pooled tests

Reliability parameter	Optimal test		Suboptimal test	
	Estimate	95% CI	Estimate	95% CI
1-PSp	5%	1 to 15%	4%	1 to 9%
$PSe_{0.02}$	92%	75 to 99%	8%	4 to 16%
$PSe_{0.80}$	>99%[a]	>99%	48%	42 to 53%

[a] Values denoted >99% are ≥99.5%

The plots (Fig. 3) illustrate how the probability of a positive test result for a pool changes as the expected number of infected individuals in a pool increases. Summaries of predicted values (with 95 percent confidence intervals [CI] of predictions) for 1-PSp, $PSe_{0.02}$ and $PSe_{0.80}$ are also given (Table 1). These results are for a single simulated dataset. Deviations between predicted values obtained from the model regression parameters and the observed values obtained directly from 50 simulated datasets will detect bias to the predicted values. Results for the optimal test (Table 2) demonstrate that predicted values of PSp are substantial underestimates of the true values, and that this substantial bias also affects predicted values of PSe for pools containing one infected individual which are also underestimated. There is some evidence of a minor underestimation of PSe for pools containing two infected individuals on some occasions. In contrast there is little evidence for a consistent bias for the suboptimal test where deviations are more evenly distributed on either side of the observed values.

Table 2. Deviations between predicted and observed values

Number of infected individuals in pool	Optimal test		Suboptimal test	
	Median	95% PCTL[a]	Median	95% PCTL
0	6%[b]	2 to 12%	1%	-4 to 7%
1	-6%	-20 to 2%	1%	-11 to 10%
2	-1%	-3 to 0%	1%	-15 to 14%
3	0%	0 to 0%	4%	-18 to 18%
4	0%	0 to 0%	2%	-15 to 16%
5	0%	0 to 0%	0%	-9 to 8%

[a] 95 percentiles of deviations; [b] deviations are predicted-observed value

RESULTS FOR A REAL INFECTION

Pancreas Disease (PD) is an infectious (Raynard & Houghton, 1993) viral (Nelson et al., 1995) disorder affecting cultivated Atlantic salmon, *Salmo salar* L. (Munro et al., 1984). The condition has been associated with high mortality (Wheatley et al., 1995), reduced growth (Munro et al., 1984) and reduced flesh quality at harvest (Lerfall et al., 2012) leading to substantial direct costs (Aunsmo et al., 2010). The virus has been classified as being a species of the genus *Alphavirus* (Weston et al., 1999) and named *Salmonid alphavirus* (SAV) (Weston et al., 2002). At the time of writing the OIE is considering whether to list infection with SAV. Diagnostic tests used to detect SAV infected individuals include virus isolation by cell culture (e.g. Nelson et al., 1995) and reverse transcriptase PCR-based technology (e.g.

Hodneland & Endresen, 2006). To date there are no published estimates of the reliability of tests for SAV based on these technologies under field conditions for either individuals or pools.

Tissue samples were taken from cultivated Atlantic salmon, *Salmo salar* L., on a marine farm experiencing an outbreak of PD on the northwest coast of Scotland during 2008. A total of 600 fish, each of mass 2.2-7.0 kg, were sampled without conscious selection from four groups comprising two pens at two time points a month apart. Tissue samples taken comprised combined head, kidney and heart ventricle stored in a viral transport medium at approximately 4 ºC for up to two days and then at -20 ºC until testing by virus isolation and kidney stored in RNALater at approximately 4 ºC until testing by qRT-PCR. Virus isolation involved the culture of SAV in Chinook salmon embryo (CHSE-214) cells from the supernatant of homogenised tissue following neutralisation for infectious pancreatic necrosis virus. Potentially infected cultures were identified from a cytopathic effect (CPE) and confirmed using qRT-PCR. The qRT-PCR test involved RNA extraction from either homogenised tissue (for the field samples) or cells (for cultures with CPE), transcription of RNA to cDNA, and qPCR amplification and detection of the 5' region of the SAV nsP1 gene (Hodneland & Endresen, 2006). Pools comprising the tissue of five individuals were assembled following tissue homogenisation and subjected to the same laboratory tests as the individuals. Estimates of Se and Sp for the individual tests and the prevalence of SAV-infected fish in each group were calculated from the posterior distribution of multiple MCMC chains generated by slice sampling assuming conditional independence (Branscum et al., 2005). Priors comprised the probability density distributions U(0,1) for Se and prevalence, and U(0.5-1) for Sp. Possible conditional dependence between the tests was not modelled because there were insufficient degrees of freedom available. The individual estimates were then used to model the reliability of the pooled test by regression as described in the previous sections of this report.

There were 94 virus isolation and 73 qRT-PCR positive individuals. Estimates of Sp were greater than 99 percent for both the virus isolation and qRT-PCR tests. Estimates of Se for the virus isolation and qRT-PCR tests were 50 percent (with 95 percent CI of 41-62 percent) and 39 percent (31-49 percent) respectively. The difference in Se is unlikely to be a consequence of chance as assessed using the Bayesian posterior probability ($p=0.013$). Estimates of the prevalence of SAV infected individuals for each sampled group are 0.37 (0.26-0.50) and 0.01 (<0.01-0.04) for the first and second time points for pen 1 and 0.86 (0.71-0.99) and 0.02 (<0.01-0.05) for the first and second time points for pen 2.

There were 34 virus isolation positive pools and 24 qRT-PCR positive pools. Thirty-three of these positive pools contained one or more virus isolation and/or qRT-PCR positive individual(s) and seven contained no positive individuals. A more detailed summary of the observed results is presented in Table 3.

The predicted results for detecting infection with SAV are presented below. The plots (Fig. 4) illustrate how the predicted probability of a positive pooled virus isolation and qRT-PCR result change as the number of infected individuals in a pool increases.

Table 3. Observed test status of pools by number of test positive individuals in pools

Number of positive individuals in pool[a]	Virus isolation		qRT-PCR	
	Positive	Negative	Positive	Negative
0/5	3	71	4	71
1/5	10	5	5	11
2/5	6	5	11	10
3/5	9	1	3	0
4/5	5	0	1	0
5/5	1	0	0	0

[a] For individuals testing either virus isolation and/or qRT-PCR positive

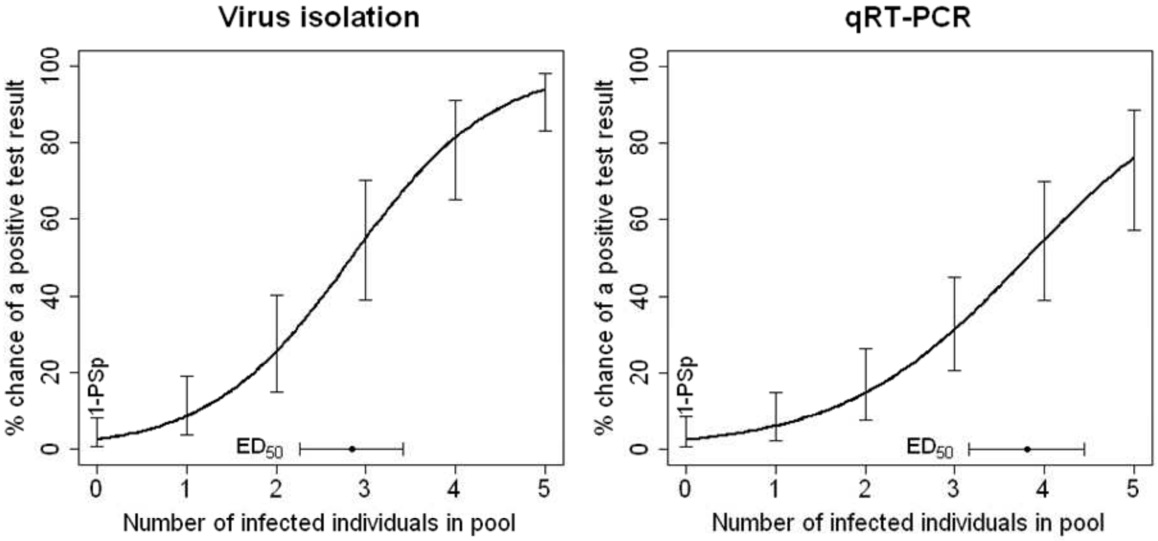


Fig. 4 Plots of predicted reliability for pooled tests detecting SAV

The results presented above estimate PSp for the virus isolation and qRT-PCR tests to be 97 and 98 percent, respectively. Information from the theoretical infection suggests that these may be underestimates of the true values. An independent estimate of PSp can be obtained by carrying out the standard Bayesian estimation of sensitivity and specificity using MCMC with pooled rather than individual test results. Values of the prevalence of infected individuals in each group obtained from the individual analysis can be used to estimate the prevalence of pools containing one or more infected individuals reducing the need to infer these parameters and increasing the chance that the model, which is based on only a fifth of the number of observations used for the individual analysis, will be identifiable. The results of such an analysis, carried out on these data, generate estimates of PSp of 99 percent (with CI of 94 to >99 percent) for both pooled tests. Estimates of PSp using this approach are to be preferred to those from the regression intercept. A summary of the reliability estimates for the SAV tests is presented below (Table 4).

Table 4. Summary of predicted reliability parameters for pooled tests detecting SAV

Reliability parameter	Virus isolation		qRT-PCR	
	Estimate	95% CI	Estimate	95% CI
1-PSp	1%	<1 to 6%	1%	<1 to 6%
$PSe_{0.02}$	9%	4 to 20%	6%	3 to 18%
$PSe_{0.80}$	76%	68 to 82%	54%	48 to 62%

DISCUSSION

The utility of pooled samples in a veterinary context has been investigated for over 20 years (Sacks et al., 1989). Previous experimental evaluations of pooled tests have involved estimating the proportion of test positive pools containing putative infected material comprising tissues spiked with laboratory cultured strains, tissues from artificially infected experimental animals or different numbers of individual test positive field samples. While these studies have provided useful information they have not generated estimates of standardised test reliability parameters, similar to Se and Sp for individual tests, required by the OIE (World Organisation for Animal Health, 2012). This report has described an approach which provides such estimates.

The main problem with the approach described in this report is bias for some tests in estimates of PSp and also PSe for pools containing one infected individual. Attempts to resolve this problem within the context of the maximum likelihood approach used for this investigation by weighting pools by the inverse of the variance of the expected number of infected individuals, using alternative link functions, and the inclusion of a quadratic explanatory variable have failed to resolve the problem satisfactorily. Likewise preliminary attempts to model the expected variation of the number of infected individuals in a pool within the context of MCMC have not, to date, resolved the problem. The solution adopted in this report is to estimate PSp separately using the non-gold standard MCMC approach to estimating sensitivity and specificity (Branscum et al., 2005) using pooled rather than individual test results. This is not always feasible because there are a smaller number of pools than individuals although prior knowledge of the prevalence of infected individuals in each population does provide information on the expected proportions of infected pools which can assist in generating an identifiable model. This does not, however, overcome the potential underestimation of PSe for pools containing one infected individual. It is recommended, at the present time, that estimates of PSp and PSe for pools containing one infected individual obtained using regression be regarded as minimum estimates of the true values.

Despite the problems described in the previous paragraph the approach outlined in this report does generate useful information which allows a comparison of individual and pooled diagnostic tests for SAV. The individual virus isolation and qRT-PCR tests can both be regarded as suboptimal with Se of less than 80 percent. This may be a consequence of either the technology upon which the tests are based or the particular implementation of the technology. The suboptimal Se of the qRT-PCR test is not consistent with laboratory experience indicating that the primers, probe and reaction conditions are likely to be suitable for detecting SAV infected individuals in the field (Hodneland & Endersen, 2006; Graham et al., 2011). Rather it is more likely that the suboptimal Se is a consequence of implementation. A review of the distribution and burden of SAV in the tissues of infected individuals (e.g. Graham et al., 2010) suggests that the inclusion of heart as well as kidney may contribute additional SAV RNA which could improve the reliability of the test. In contrast the main constraint to reliable virus isolation is likely to be the technology used. Successful virus isolation requires that the SAV within samples is viable on receipt, is efficiently replicated in the cell line used and, for the test described in this report, causes CPE. The CPE associated with SAV can be indistinct (Desvignes et al., 2002) and it is possible that the adoption of immunostaining to detect SAV-infected cells (Graham et al., 2003) would improve Se.

The pooled virus isolation and qRT-PCR tests are similarly suboptimal with PSe of less than 80 percent. The estimates of $PSe_{0.02}$ for the virus isolation and qRT-PCR tests are approximately nine and six percent, respectively, and, while it is possible that these are subject to a small amount of underestimation, the values suggest that the use of individuals rather pools is more appropriate when testing for freedom from disease. Estimates of $PSe_{0.80}$ are substantially higher (approximately 76 and 54 percent for virus isolation and qRT-PCR respectively) and initially suggest that the pooled tests can be used for the confirmatory diagnosis of a suspected infection if two or three pools are used. Estimates of PSp of 99 percent for both the virus isolation and qRT-PCR pooled tests are comparable with the compound Sp of tests on five individuals. Unfortunately the necessity of using two or three pools to reliably confirm suspected infection does reduce the compound PSp for multiple pools to below the compound Sp for five individuals. Overall, the results therefore suggest that the testing of individuals is also to be preferred for the confirmatory diagnosis of a suspected SAV infection.

The inability to incorporate conditional dependence into the models for SAV because of a lack of a sufficient number of degrees of freedom is unfortunate because the virus isolation test uses qRT-PCR to confirm the presence of SAV in cultures showing CPE. Failure to incorporate conditional dependence has been demonstrated to have an effect on individual test reliability and prevalence estimates (Vacek, 1985) and can, therefore, affect pooled test reliability estimates. While the actual values of the SAV test reliability parameters are subject to an unknown amount of bias it is unlikely that conditional dependence will have a substantive effect on the conclusions above because the estimates of Se and PSe are both clearly suboptimal.

This report has described an approach to evaluating the reliability of diagnostic tests on pooled groups of individuals which generates test reliability parameters consistent with OIE requirements. While the approach has limitations it has provided useful information on the reliability of a pooled test for SAV and can be refined and superseded over time.

ACKNOWLEDGEMENTS

We thank the operators and staff of the aquaculture business for their assistance with the collection of samples. Sampling, laboratory testing, statistical analysis and discussion have involved many Marine Scotland scientists, including Anna Donald, John McLean, Nichola McManus, Mark Fordyce, Mickael Fourrier, Rachel Kilburn, Katherine Lester, Eann Munro, Warren Murray, Campbell Pert and Sarah Weir whose help is gratefully acknowledged. This work was funded by The Scottish Government.

REFERENCES

Aunsmo, A., Valle, P.S., Sandberg, M., Midtlyng, P.J. and Bruheim, T. (2010). Stochastic modelling of direct costs of pancreas disease (PD) in Norwegian farmed Atlantic salmon (*Salmo salar* L.). Prev. Vet. Med. 93, 233-241

Branscum, A.J., Gardner, I.A. and Johnson, W.O. (2005). Estimation of diagnostic-test sensitivity and specificity through Bayesian modeling. Prev. Vet. Med. 68, 145-163

Dendukuri, N. and Joseph, L. (2001). Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests. Biometrics 57, 158-167

Desvignes, L., Quentel, C., Lamour, F. and Le Ven, A. (2002). Pathogenesis and immune response in Atlantic salmon (Salmo salar L.) parr experimentally infected with salmon pancreas disease virus (SPDV). Fish Shellfish Immunol. 12, 77-95

Dorfman, R. (1943). The detection of defective members of large populations. Ann. Math. Stat. 14, 436-440

Enøe, C., Georgiadis, M.P. and Johnson, W.O. (2000). Estimation of sensitivity and specificity of diagnostic tests and disease prevalence when the true disease state is unknown. Prev. Vet. Med. 45, 61-81

Firth, D. (1993). Bias reduction of maximum likelihood estimates. Biometrika 80, 27-38

Georgiadis, M.P., Gardner, I.A. and Hedrick, R.P. (1998). Field evaluation of sensitivity and specificity of a polymerase chain reaction (PCR) for detection of *Nucleospora salmonis* in rainbow trout. J. Aquat. Anim. Health 10, 372-380

Graham, D.A., Fringuelli, E., Wilson, C., Rowley, H.M., Brown, A., Rodger, H., McLoughlin, M.F., McManus, C., Casey, E., McCarthy, L.J. and Ruane, N.M. (2010). Prospective longitudinal studies of salmonid alphavirus infections on two Atlantic salmon farms in Ireland; evidence for viral persistence. J. Fish Dis. 33, 123-135

Graham, D.A., Frost, P., McLaughlin, K., Rowley, H.M., Gabestad, I., Gordon, A. and McLoughlin, M.F. (2011). A comparative study of marine salmonid alphavirus subtypes 1-6 using an experimental cohabitation challenge model. J. Fish Dis. 34, 273-286

Graham, D.A., Jewhurst, V.A., Rowley, H.M., McLoughlin, M.F. and Todd, D. (2003). A rapid immunoperoxidase-based virus neutralization assay for salmonid alphavirus used for a serological survey in Northern Ireland. J. Fish Dis. 26, 407-413

Hodneland, K. and Endresen, C. (2006). Sensitive and specific detection of *Salmonid alphavirus* using real-time PCR (TaqMan®). J. Virol. Methods 131, 184-192

Hui, S.L. and Walter, S.D. (1980). Estimating error rates of diagnostic tests. Biometrics 36, 167-171

Kennedy, J.A. (2006). Diagnostic efficacy of a reverse transcriptase-polymerase chain reaction assay to screen cattle for persistent bovine viral diarrhea virus infection. J. Am. Med. Vet. Assoc. 229, 1472-1474

Kinde, H., Read, D.H., Chin, R.P., Bickford, A.A., Walker, R.L., Ardans, A., Breitmeyer, R.E., Willoughby, D., Little, H.E., Kerr, D. and Gardner, I.A. (1996). *Salmonella enteritidis*, phage type 4 infection in a commercial layer flock in southern California: bacteriologic and epidemiologic findings. Avian Dis. 40, 665-671

Lerfall, J., Larsson, T., Sveinung, B., Taksdal, T., Dalgaard, P., Afanasyev, S., Bjerke, M.T. and Mørkøre, T. (2012). Effect of pancreas disease (PD) on quality attributes of raw and smoked fillets of Atlantic salmon (*Salmo salar* L.). Aquaculture 324-325, 209-217

McBeath, A.J.A., Bain, N. and Snow, M. (2009). Surveillance for infectious salmon anaemia virus HPR0 in marine Atlantic salmon farms across Scotland. Dis. Aquat. Org. 87, 161-169

Munro, A.L.S., Ellis, A.E., McVicar, A.H., McLay, H.A. and Needham, E.A. (1984). An exocrine pancreas disease of farmed Atlantic salmon in Scotland. Helgol. Meeresunters. 37, 571-586

Neal, R.M. (1997). Markov chain monte carlo methods based on ‘slicing’ the density function. Technical Report 9722, Department of Statistics, University of Toronto. Obtained from www.cs.utoronto.ca/~radford/publications.html on 5th December 2012

Nelson, R.T., McLoughlin, M.F., Rowley, H.M., Platten, M.A. and McCormick, J.I. (1995). Isolation of a toga-like virus from farmed Atlantic salmon *Salmo salar* with pancreas disease. Dis. Aquat. Org. 22, 25-32

Nérette P., Dohoo I. and Hammell L. (2005). Estimation of sensitivity and specificity of three diagnostic tests for infectious salmon anaemia virus in the absence of a gold standard. J. Fish Dis. 28, 89-99

Raizman, E.A., Scott, J.W., Muñoz-Zanzi, C.A. and Tavornpanich, S. (2011). Estimated within-herd prevalence (WHP) of *Mycobacterium avium* subsp. *paratuberculosis* in a sample of Minnesota dairy herds using bacterial culture of pooled fecal samples. Can. J. Vet. Res. 75, 112-116

Raynard, R.S. and Houghton, G. (1993). Development towards an experimental protocol for the transmission of pancreas disease of Atlantic salmon *Salmo salar*. Dis. Aquat. Org. 15, 123-128

Sacks, J.M., Bolin, S.R. and Crowder, S.V. (1989). Prevalence estimation from pooled samples. Am. J. Vet. Res. 50, 205-206

Vacek, P.M. (1985). The effect of conditional dependence on the evaluation of diagnostic tests. Biometrics 41, 959-968

Weston, J., Villoing, S., Brémont, M., Castric, J., Pfeffer, M., Jewhurst, V., McLoughlin, M., Rødseth, O., Christie, K.E., Koumans, J. and Todd, D. (2002). Comparison of two aquatic alphaviruses, salmon pancreas disease virus and sleeping disease virus, by using genome sequence analysis, monoclonal reactivity, and cross-infection. J. Virol. 76, 6155-6163

Weston, J.H., Welsh, M.D., McLoughlin, M.F. and Todd, D. (1999). Salmon pancreas disease virus, an alphavirus infecting farmed Atlantic salmon, *Salmo salar* L. Virology 256, 188-195

Wheatley, S.B., McLoughlin, M.F., Menzies, F.D. and Goodall, E.A. (1995). Site management factors influencing mortality rates in Atlantic salmon (*Salmo salar* L.) during marine production. Aquaculture 136, 195-207

World Organisation for Animal Health (2009). Manual of Diagnostic Tests for Aquatic Animals. World Organisation for Animal Health, Paris, 6th edition

World Organisation for Animal Health (2012). Aquatic Animal Health Code. World Organisation for Animal Health, Paris, 15th edition

BAYESIAN NETWORK MODELLING OF BIOSECURITY PERCEPTIONS, PRACTICE AND INFECTION IN THE 2007 EQUINE INFLUENZA OUTBREAK IN AUSTRALIA

S.M. FIRESTONE*, K. SCHEMANN, F.I. LEWIS, M.P. WARD, J.-A.L.M.L. TORIBIO,
M.R. TAYLOR AND N.K. DHAND

SUMMARY

Australia experienced its first ever outbreak of equine influenza in August 2007. Nearly 70,000 horses were infected over a period of 5 months before the disease was successfully eradicated through the combination of horse movement controls, on-farm biosecurity and vaccination. In a previous premises-level case-control study of the 2007 equine influenza outbreak in Australia, the protective effect of several variables representing on-farm biosecurity practices were identified. Separately, factors associated with horse managers' perceptions of the effectiveness of biosecurity measures have been identified.

In this analysis additive Bayesian network modelling is applied to describe the complex web of associations linking variables representing on-farm human behaviours during the 2007 equine influenza outbreak (compliance or lack thereof with advised personal biosecurity measures) and horse managers' perceptions of the effectiveness of such measures in the event of a subsequent outbreak. It was observed that past performance in an outbreak setting may be a reliable predictor of future perceptions, and should be considered when targeting infection control guidance to horse owners and managers.

INTRODUCTION

Equine influenza is a major cause of equine respiratory disease worldwide (Daly et al., 2004). Although most horses recover uneventfully, large rapidly-spreading outbreaks have occurred when highly susceptible horse populations have been exposed to novel virus strains (Uppal & Yadav, 1987; Dalglish, 1992; Guthrie et al., 1999). Such outbreaks can cause major disruption to horse racing and other equestrian events (Daly et al., 2006) leading to severe social and economic impacts (Callinan, 2008) and psychological distress (Taylor et al., 2008) amongst horse owners and other people involved in the horse industry.

Prior to August 2007, Australia was one of only three countries reported to be free of equine influenza virus (OIE, 2012). After the introduction of subclinically infected horses from Japan and release from a quarantine station in Sydney, infection disseminated through transmission via several equestrian events and the movement of infected horses prior to disease detection on 24 August 2007 (Callinan, 2008; Watson et al., 2011; Firestone et al.,

---

* Simon Firestone, Faculty of Veterinary Science, The University of Melbourne, Parkville, VIC 3010, Australia. Email: Simon.Firestone@unimelb.edu.au

2012). Around 70,000 horses on 9,359 premises were infected with equine influenza A/H3N8 virus over a period of 5 months before the disease was successfully eradicated through the combination of zone-based horse movement controls, a public awareness programme that included outbreak communications targeting horse owners and the horse industry, providing updates and suggesting on-farm biosecurity measures to reduce local spread of disease, and an emergency vaccination campaign (Equine Influenza Epidemiology Support Group, 2008).

Compliance with certain on-farm biosecurity practices (particularly having a footbath in place before any horses were infected) was associated with a four-fold reduction in odds of infection in a previous retrospective case-control study of the 2007 equine influenza outbreak in Australia (Firestone et al., 2011). Personal hygiene practices of visitors (changing shoes and clothes, washing hands and on some premises showering before contact with uninfected horses) that had contact with the horses on a premises, and the regularity of horse handling were observed to be important factors in determining whether or not horses on a premises were infected (Firestone et al., accepted for publication). Other risk factors included the proximity to nearby infected premises and horse attendance at events where transmission was known to have occurred.

Social cognition theories, such as the protection motivation theory (Maddux & Rogers, 1983; Norman et al., 2005), were originally developed to explore health behaviours. More recently these theories have been successfully applied to horse owner's biosecurity practices (Schemann et al., 2011). Such theories consider the relative contributions of an individual's threat appraisal and coping appraisal on the motivation to take-up or comply with recommended protective behaviours; in this current context, on-farm biosecurity practices. This approach has been described previously (Schemann et al., 2012a). In the health behaviour literature there is limited research on the relationships between prior health behaviours and risk exposure/outcomes and individual's perceptions of the effectiveness of those behaviours and their future preparedness in relation to a future health threat (Taylor et al., 2012). It is generally acknowledged that prior experience impacts on threat and coping appraisals and prior behaviour is a good predictor of future behaviour (Umeh, 2004). Similarly one would expect that a positive association between a protective practice and a positive threat outcome (e.g. avoidance of infection) would result in a more positive coping appraisal in relation to that same threat in the future; one in which an individual would feel better able to manage a future threat (increased self-efficacy) and would regard the practice as effective against a future threat (high response efficacy) and would therefore be more likely to execute that behaviour in the event of that threat reoccurring in the future.

Bayesian network (BN) modelling (Buntine, 1991; Heckerman et al., 1995) is a modern graphical approach for analysing complex systems of interdependent variables. In contrast to other multivariate techniques (such as factor analysis and principal components analysis), graphical modelling does not involve any dimension reduction, instead retaining as many characteristics of the underlying system as possible whilst facilitating interpretation. The challenge is to devise such a graphical structure. In BN modelling, an optimal graphical model is derived in a highly objective manner directly from empirical data and depicted as a directed acyclic graph (DAG) (Lewis et al., 2011). Such objectivity comes at a cost; BN analyses are largely exploratory: they provide evidence of statistical dependencies only, rather than being causal models. In this respect they are comparable to traditional statistical modelling approaches, such as linear regression models. The defining difference is that a BN analysis determines a multidimensional model of the observed data with all variables potentially dependent on each other (Lewis & McCormick, 2012).

Previous separate analyses have been conducted into risk factors (Firestone et al., 2011) and horse owner and manager perceptions (Schemann et al., 2012a; Schemann et al., 2012b) based on interview data collected during a case-control study into the 2007 equine influenza outbreak. Exact additive Bayesian network (ABN) modelling has also been applied to delve into correlations and inter-dependencies in these data (Firestone et al., accepted for publication; Schemann et al., accepted for publication), again separately for biosecurity perceptions and practices. The current, larger ABN analysis bridges the gap by seeking to identify linkages between perceptions and practice. The aim of this analysis was to describe the complex web of associations between on-farm human behaviours during the 2007 equine influenza outbreak (compliance or lack thereof with advised personal biosecurity measures) and horse managers' perceptions of the effectiveness of such measures in the event of a subsequent outbreak.

MATERIALS AND METHODS

The dataset

This analysis was based on data from interviews conducted during a retrospective case-control study of the 2007 outbreak of equine influenza in Australia (Firestone et al., 2011). The questionnaire design, sampling approaches and data management are described in detail elsewhere (Firestone et al., 2011). A sample of 270 premises was randomly selected from a dataset of all horse premises that were laboratory tested for equine influenza virus (H3N8) by the New South Wales (NSW) Department of Primary Industries, using real-time reverse transcription polymerase chain reaction (Foord et al., 2009), in the first seven weeks of the outbreak. All eligible premises were located in highly affected areas of NSW: the Restricted Area ('red zone') or Special Restricted Area ('purple zone') designated according to a risk-based zoning system (NSW DPI, 2007). The owners and managers of these horse premises were sent personally addressed letters requesting their participation in the study and up to three follow-up telephone calls were made, as approved by the Human Research Ethics Committee of the University of Sydney (07-2009/11840). Two trained interviewers conducted structured face-to-face interviews on-farm between July and October 2009. Interviews included a series of recall-improving activities and were based on a piloted questionnaire containing a total of 61 questions. All data were entered into a purpose-built Microsoft Access 2007 database (Microsoft Corporation, Redmond, WA, USA). Basic data manipulation, range checks and data cleaning were conducted on this database and in the R statistical package (R Development Core Team, 2011).

The 31 variables included in this analysis are presented in Table 1. To reduce the computational complexity of the BN modelling it was necessary to re-classify variables dichotomously, as per our previous analyses (Firestone et al., 2011; Schemann et al., 2012a). For the nine ordinal variables representing personal biosecurity practices the levels 'every time' and 'most times' were grouped, as were 'sometimes' and 'never', into nine corresponding binary variables. The variable representing horse managers' perceptions of their level of stringency in applying on-farm biosecurity measures during the entire 2007 outbreak period was re-classified into the binary variable 'Highly stringent' by grouping the lower two ordinal levels: 'normal practices' with 'less stringent than normal practices'. The 17 ordinal variables representing horse managers' perceptions of the effectiveness of certain biosecurity measures were each re-classified by grouping the levels 'partially effective' and 'not effective', thus making 17 binary variables with categories of 'very effective' and ‘partially or not effective’. The variables representing horse managers’ perceptions of their level of vulnerability to and preparedness for future outbreak were re-classified into the binary variables ‘Highly vulnerable’ and ‘Highly prepared’ by grouping the lower two ordinal levels: ‘vulnerable’ with ‘not vulnerable’, and ‘prepared’ with ‘unprepared’, respectively.

Table 1. The 31 binary variables representing infection status, on-farm biosecurity practice and perceptions of the effectiveness of biosecurity measures from interviews of 148 horse managers

Variable	Description
Infection status	
Purple zone	
Personal biosecurity practices conducted prior to contacting the horses (most or every time):	
Footbath	A footbath was in place
Hands	People from the premises washed their hands
Clothes	People from the premises changed their clothes
Shoes	People from the premises changed their shoes
Shower	People from the premises showered
Visitor hands	Visitors to the premises washed their hands
Visitor clothes	Visitors to the premises changed their clothes
Visitor shoes	Visitors to the premises changed their shoes
Visitor shower	Visitors to the premises showered
Highly stringent *(perceived)*	On-farm measures undertaken during the 2007 outbreak were more stringent than normal practices
Perceived effectiveness of on-farm biosecurity measures:	
No movement	Complying with movement restrictions
Restrict horse contact	Reducing managed horses’ contact with other horses
Isolate horses	Ensuring that any new horses are isolated from other managed horses
Disinfect vehicles	Disinfecting vehicles entering the managed premises
Disinfect floats	Disinfecting horse floats before use
Restrict own contact	Reducing own contact with other horses
Wash hands	Washing hands before contact with managed horses
Use soap	Using soap when washing hands
Change clothes	Changing into clean clothes on arrival at the managed premises
Change shoes	Changing your shoes on arrival at the managed premises
Shower	Showering on arrival at the managed premises
Control access	Controlling who has access to managed horses
Restrict professional visits	Reducing visits by horse professionals to the managed premises
Use a footbath	Ensuring that all visitors to the premises use a disinfectant footbath
Not share equipment	Not sharing horse equipment
Clean equipment	Cleaning horse equipment before use
Clean feed source	Ensuring that feed and bedding comes from a clean source
Perceived readiness to respond to a future outbreak:	
Highly prepared	Highly prepared for a future outbreak
Highly vulnerable	Highly vulnerable to a future outbreak

Additive Bayesian networks

A robust statistical model of the 31 binary variables was identified using ABN modelling, following the approach of Lewis & McCormick (2012). Determining an optimal DAG involved identifying a high scoring consensus network structure and adjusting the chosen model for over-fitting. All modelling was carried out in R, using the abn library (Lewis, 2012). Highly computational analyses were implemented in parallel across 64 CPUs on the University of Melbourne's high performance computing cluster 'edward' (each CPU having 2Gb of RAM). A heuristic search approach was used to identify highly conserved network structures (DAGs) in tens of thousands of high scoring networks by implementing a random restarts greedy hill-climber technique (Heckerman et al., 1995). This approach is not guaranteed to find the best possible model (in contrast to an exact approach) but will potentially identify many well-fitting models which can be used to produce a highly robust consensus (summary) model.

Each search started from a different randomly chosen DAG structure. At each step an arc was either added, removed or its direction reversed, and all possible moves are computed until no further improvement in model goodness of fit is possible. The (log) marginal likelihood (MacKay, 1992) was used to compare all models. Such analyses require prior distributions on both the DAG structures and the model parameters. A uniform structural prior was used in which, in the absence of any data, all DAG structures were equally plausible. Uninformative parameter priors were used throughout, specifically Gaussian distributions with means of zero and variance of 1000 for the marginal log odd ratios in each node (a node is an individual multivariable logistic regression). By using uninformative parameter priors, structure discover effectively commenced from a point equivalent to no prior knowledge (or data). Specifying and justifying informative parameter priors is impractical for every combination of variables across every DAG under comparison.

A well known issue with any automated model selection process, including common approaches such as stepwise searches with linear regression models, is a strong tendency to overfit (Babyak, 2004). The same applies to ABN modelling. This was addressed by model averaging to build a consensus network based on the initial searches (Hoeting et al., 1999) followed by parametric boot-strapping of the consensus network (Friedman et al., 1999) to further prune unsupported structural detail. To build the consensus network the hill-climber search algorithm was implemented with a typical and necessary refinement to constrain the search space (Lewis et al., 2011) that limits the number of parents per node (applying a 'parent limit'). First, 20,000 stepwise searches were conducted at a parent limit of one. This large set of locally optimal ABN models was randomly divided into two groups, and within each a directed majority consensus network built from all arcs that appeared in >50% of the locally optimal ABNs (in each group). A robust summary model was considered to have been identified (within the current parent limit) if the consensus networks from the two randomly selected groups were identical. The parent limit was incremented and the 20,000 searches repeated until the parent limit exceeded the maximum number of parent nodes of any variable in the consensus DAG. Following Lewis and McCormick (2012), the search space was further constrained for parent limits >3 by only considering additional arcs for variables which had met the parent limit in the previous step.

The stable consensus network model was then used to generate 5,000 bootstrap datasets (simulations which are each the same size as the original dataset) using the open source JAGS software (http://mcmc-jags.sourceforge.net/). A single search was conducted on each of these bootstrap samples, at each parent limit, and the same test applied for robustness of the consensus network on two randomly selected groups. To address over-fitting, any arcs in the stable consensus network from the original data that were not recovered (in either direction) in >50% of the bootstrap DAGs were deemed to have insufficient statistical support to be considered robust (Poon et al., 2007; Lewis et al., 2011; Lewis & McCormick, 2012).

The endpoint of the analyses is a statistically robust ABN with the parameters in the model being exactly those in a standard multivariable logistic regression, except there are many more of these: a set for each variable in the study data. The parameters have the usual interpretation as posterior marginal log odds ratios for each covariate. These were estimated, along with their posterior 95% confidence intervals, using the abn library which applies a standard numerical approach in Bayesian estimation – Laplace approximations (Tierney & Kadane, 1986). Arc directions were collapsed once a robust network of statistical dependencies was identified (Poon et al., 2007; Lewis & McCormick, 2012). Although arc directions are necessary for the mechanics of ABN modelling they do not imply causal directions and were not considered biologically meaningful in this analysis.

RESULTS

Of the 270 premises selected, 38 were ineligible for enrolment for various reasons, whilst the owners and managers of another 32 premises declined to participate, resulting in a study size of 200 premises. A number of questions were deemed ‘not applicable’ by horse managers’ of certain premises (for example: when asked about disinfecting a horse float, when they did not own or use a horse float); these were treated as missing data. The final dataset consisted of 148 observations.

A stable majority consensus ABN model was constructed by pooling search results across the initial 20,000 heuristic searches (at a parent limit of six). This preliminary model had 48 arcs and a log marginal likelihood of -2243.10. Parametric bootstrapping based on the directed consensus model took 15 hours runtime per block of 100 samples (with associated searches) at each parent limit. At a parent limit of six, no node had more than 5 parents and the median number of arcs in the 5,000 bootstrapped majority consensus DAGs was 39, with a maximum of 40 arcs. This is compelling evidence that the observed data do not support the level of structural complexity in the preliminary model. Pruning off arcs that were not recovered in at least 50% of the bootstrapped DAGs resulted in the removal of 9 arcs. The same bootstrapping analysis was performed, taking 100 different random subsets of 2,500 bootstrap samples; each gave identical results in terms of the arcs to be removed, which is strong evidence that sufficient bootstrap samples were generated.

The final undirected majority consensus DAG with 39 arcs is presented in Fig. 1. This Bayesian network model was constructed at a parent limit of six and had a log marginal likelihood of -2364.06. The final DAG model comprised three subsets of nodes connected to each other but disconnected from all other nodes (‘components’). The largest of these included all of the on-farm personal biosecurity measures studied and most of the perceptions. However these were tightly clustered into two distinct sub-groups connected by only three ‘bridging’ links.

The arcs in Fig. 1 are labelled with estimates of the posterior median odds ratios. Personal biosecurity practices are represented by grey circles, horse managers' perceptions of the effectiveness of biosecurity measures in a future outbreak are represented by white squares, whilst other perception variables are represented by black hexagons. The premises zone and infection status at the end of the 2007 outbreak are represented by triangles, and the grey dotted arc represents a negative association.

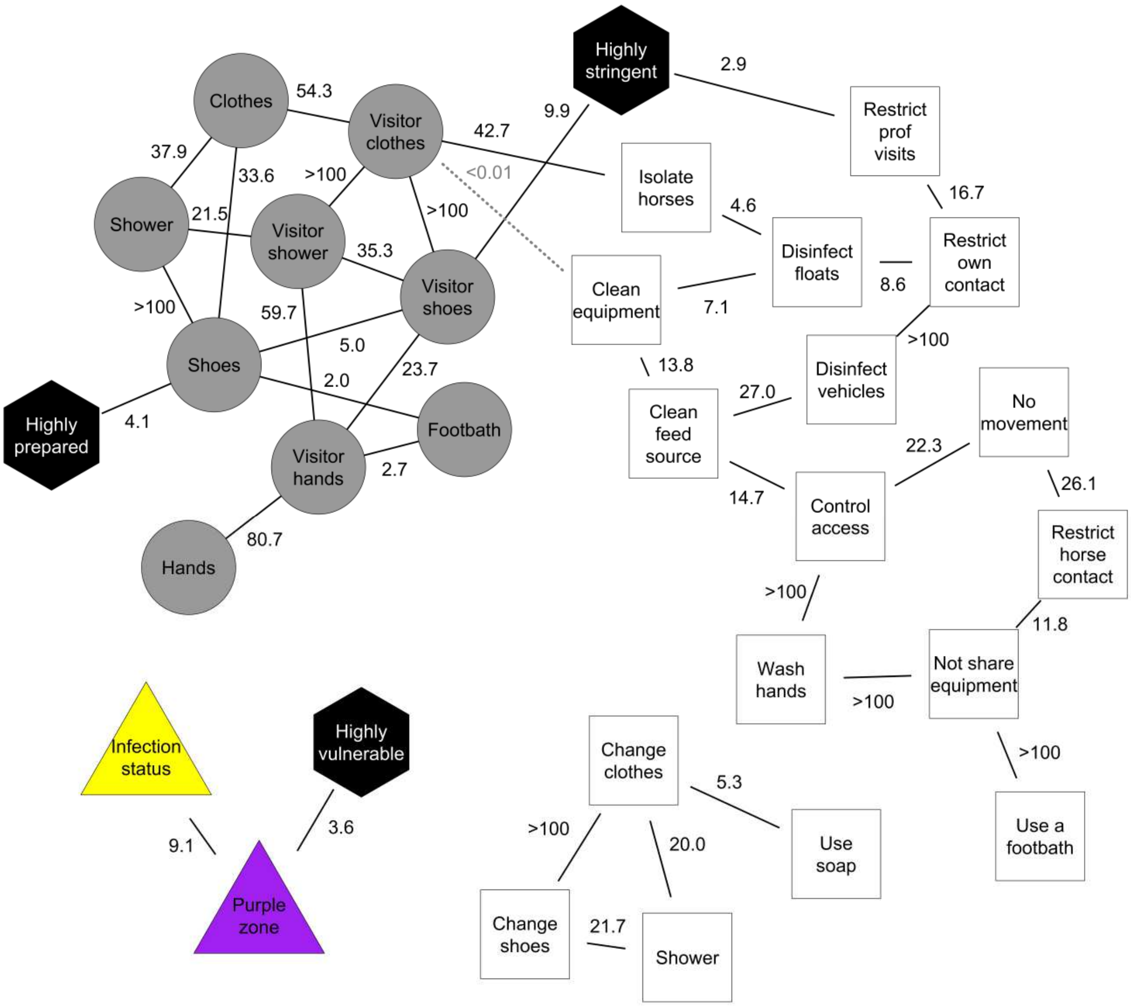


Fig. 1 Majority consensus Bayesian network model (after bootstrapping to prune structural detail unsupported by the data, n=148) of the inter-relationship between personal biosecurity practices in the 2007 equine influenza outbreak in Australia and horse managers' perceptions of the effectiveness of certain measures in a future outbreak

DISCUSSION

This Bayesian network analysis presents the statistical associations between horse managers' on-farm biosecurity practices during an at risk period in the 2007 outbreak with their perceptions of whether such measures will be effective in future outbreaks. Based on the final majority consensus model it appears that past biosecurity practices in an outbreak setting were directly associated with future perceptions of preparedness and the effectiveness of such measures. This is important for targeting infection control guidance to horse owners and managers.

Parts of this dataset have previously been analysed to estimate the effects of risk factors (including non-compliance with biosecurity practice) early in the outbreak (Firestone et al., 2011), and, separately, described horse manager's perceptions about the effectiveness of biosecurity practices and associated factors (Schemann et al., 2012a). Based on the findings of earlier BN analyses, on-farm biosecurity practices would be expected to be highly interlinked (Firestone et al., accepted for publication), and, respectively, perceptions of the effectiveness of on-farm biosecurity practices to be closely associated with one another (Schemann et al., accepted for publication). Many of the links are highly intuitive within these sub-groups and these have been previously described in considerable detail.

Focussing on the bridging links between the groups of variables representing perceptions and practices, the variable representing the practice of ensuring that visitors to a premises changed their clothes prior to contacting the horses (during the 2007 outbreak) was statistically associated with a perception that isolating new horses on arrival would be highly effective in preventing spread in a future outbreak (posterior median OR=42.7, Bayesian 95% CI: 2.4, 2691). Both of these variables relate to barrier infection control (on-farm and at the farm gate) because effectively quarantining horses newly introduced to a premises requires those tending and moving between the isolated horses and other groups of horses on a premises to practice stringent barrier hygiene. The personal hygiene practices of visitors (changing shoes and clothes, washing hands and on some premises showering before contact with uninfected horses) have previously been identified as some of the most important measures for determining whether or not horses on the premises were infected (Firestone et al., accepted for publication). Presumably, horse owners who were stringent in ensuring that visitors did not enter with contaminated fomites also consider it possible to keep their herd free from such a highly contagious disease by isolating new arrivals.

An association has previously been reported between greater levels of preparedness for a future outbreak and a better perception of the effectiveness of on-farm biosecurity measures (Schemann et al., 2012a). Interestingly, in the present analysis this link appears indirect. Perceived future preparedness appeared most closely associated with actual practices in the 2007 outbreak, rather than directly linked to perceptions of the effectiveness of biosecurity measures. In particular, those horse owners and managers that considered themselves 'highly prepared' were more likely to also have been stringent during the 2007 outbreak in changing their shoes before contact with uninfected horses (posterior median OR=4.1, Bayesian 95% CI: 1.2, 14.1).

The perception of highly stringent biosecurity practice (during the outbreak period) was statistically associated with ensuring visitors changed into a separate or clean pair of shoes before contacting horses (posterior median OR=9.9, Bayesian 95% CI: 3.4, 40.4) and indirectly associated with each of the other personal hygiene practices of visitors. Perceived stringency (in actual practice in 2007) was also related to a heightened perception of the effectiveness of other barrier control measures in future outbreaks, such as reducing visits by horse professionals to the managed premises, and indirectly to perceived effectiveness of horse owners and managers reducing their own contact with other horses (off the premises).

The findings of this analysis are consistent with the protection motivation theory (Maddux & Rogers, 1983; Norman et al., 2005). The observed relationships between heightened stringency and biosecurity effectiveness perceptions, and between heightened preparedness and performance of certain practices, represent a heightened coping appraisal of horse managers following the 2007 outbreak. The coping appraisal focuses on protective actions an individual can take: this depends on an individual's belief that a certain action or behaviour will reduce the threat ('response efficacy') and that he/she considers himself/herself capable of carrying out the protective measure ('self-efficacy') (Norman et al., 2005). These findings suggest that past experience of the 2007 outbreak has contributed to a heightened sense of 'response efficacy' as measured through perceived biosecurity effectiveness and of 'self-efficacy' as measured through perceived preparedness.

The third 'bridge' between perceptions and past practices was a negative association. Those horse owners and managers that ensured visitors changed their clothes before contacting horses did not consider cleaning horse equipment to be highly effective (posterior median OR=0.003, Bayesian 95% CI: 0, 0.113). A number of respondents commented at interview that as they did not share horse equipment with other premises, they did not consider cleaning equipment to be related to between premises spread. This association may therefore be explained by those who practiced stringent on-farm biosecurity either not sharing equipment between premises or not considering such a practice a major risk of infection to their horses.

A smaller network component linked infection status to the most highly affected restricted area (the 'purple zone') and perceptions of increased vulnerability. Infection status was included in this analysis as a variable representing whether or not horses on a premises were infected over the 5-month outbreak period. This variable was not expected to be associated with any of the on-farm biosecurity practice variables because these were only measured for one specific week in the early epidemic period, either prior to infection or testing. This is distinct from our previous risk factor analysis that focussed only on infection in the first 2 months of the epidemic (Firestone et al., 2011), and identified a direct association with variables representing on-farm biosecurity practice.

An important consideration when interpreting the findings of this analysis is that all questions were asked at the same face-to-face interviews in 2009, 2 years after the actual outbreak. Several aspects of the interviews were intentionally designed to minimise the potential for selection bias, poor recall, misclassification and other inaccuracies that might influence the results (Firestone et al., 2011). Importantly, questions about practices undertaken during the outbreak were anchored on a specific week of interest when the premises was tested or immediately prior to the earliest onset of infection on the premises, whereas questions about the horse managers' perceptions of the effectiveness of biosecurity measures were asked in the context of a future outbreak. Ideally this time sequence would be reversed, so one could test how perceptions influence behaviour in an ensuing outbreak. Any claims of causation should not be made in the interpretation of the findings of this analysis, more so because ABN modelling is considered an exploratory analytical ('data-mining') technique.

The present analysis was also limited in that not all biosecurity perceptions questions were applicable to all respondents, reducing the effective (complete case) study size to n=148, even though 200 premises were interviewed. Such a step introduces a selection bias and the estimates produced must always be interpreted with caution as only a single realisation of the true population estimates. Other, less desirable, options for handling these data would be to discard a number of variables (to retain the complete case size) however such an approach is prone to misspecification bias, or multiple imputation but this leads to modelling based on modelled data.

ABN modelling, as done here, or another multivariate approach, is conceptually preferable to ‘traditional’ multivariable regression-based analyses when adjusting for confounding effects and handling strongly non-independent explanatory variables (Hernan et al., 2002; Lewis et al., 2011). Applying a more general joint probability model to account for non-independence between variables may produce different results to logistic or another multivariable regression-based approach (Lewis et al., 2011; Lewis & McCormick, 2012). This is not unexpected, noting the Yule-Simpson paradox (Hand et al., 1997) and that multivariable regression-based methods may be considered a one dimensional special case of ABN modelling. At present the key limitation to ABN modelling is computational tractability. These analyses are not always practical for datasets with large numbers of variables. Indeed, in this analysis the search space had to be limited to ensure computations could be completed within a reasonable period of time on a parallel computing cluster. Although an exact structure discovery approach will identify the single best (albeit over-fitted) model for the data, currently this is only feasible for around 20 variables (Lewis, 2012). Here a heuristic search approach identified a robust statistical model of the data based on a majority consensus of a large number of high scoring networks. This approach may not identify the single best model. Although the final DAG structure may differ slightly from that resulting from an exact approach, when applied with model-averaging over sufficient searches, and bootstrapped, there are unlikely to be major interpretative differences.

In conclusion, several statistical associations were identified between horse managers’ on-farm biosecurity practices during an at-risk period in the 2007 equine influenza outbreak and their perceptions of whether such measures will be effective in a future outbreak. The practice of barrier infection control measures (on-farm and at the farm gate) was associated with a heightened perception of preparedness, whereas perceived stringency in actual practice in 2007 was related to more favourable consideration of the effectiveness of other barrier control measures. Past performance in an outbreak setting may indeed be a reliable predictor of future perceptions of preparedness and the effectiveness of on-farm biosecurity measures. These findings should be considered when targeting infection control guidance to horse owners and managers.

ACKNOWLEDGEMENTS

This research was jointly funded by the Rural Industries Research and Development Corporation (RIRDC), the Australian Biosecurity Cooperative Research Centre for Emerging Infectious Diseases (ABCRC) and the Swiss National Science Foundation (grant: IZK0Z3_140188). The authors are grateful for access to the ‘edward’ high performance computing cluster at the University of Melbourne and systems administrative support, and also for the time and cooperation of the horse premises owners and managers interviewed, the NSW DPI for making the equine influenza dataset available, and the following individuals for contributions to data compilation: Barbara Moloney (NSW DPI), Nina Kung (QLD DPIF), Brendan Cowled (AusVet) and Graeme Garner (DAFF).

REFERENCES

Babyak, M.A. (2004). What you see may not be what you get: A brief, nontechnical introduction to overfitting in regression-type models. Psychosom. Med. 66, 411-421

Buntine, W. (1991). Theory refinement on Bayesian networks. In: Kaufmann, M. (Ed.), Proceedings of Seventh Conference on Uncertainty in Artificial Intelligence, Los Angeles, CA, USA, 52-60

Callinan, I. (2008). Equine influenza - the August 2007 outbreak in Australia. Report of the Equine Influenza Inquiry. The Hon. Ian Callinan AC

Dalglish, R.A. (1992). The international movement of horses – the current infectious disease situation. In: Short, C.R. (Ed.), Proceedings of the 9th International Conference of Racing Analysts and Veterinarians, Louisiana State University, New Orleans, 37-53

Daly, J., Newton, J.R., Smith, K.C. and Mumford, J. (2006). Epidemiology of equine influenza viruses: pathogenicity and transmissibility. Rad 496. Medicinske znanosti (Croatian translated) 30, 87-94

Daly, J.M., Newton, J.R. and Mumford, J.A. (2004). Current perspectives on control of equine influenza. Vet. Res. 35, 411-423

Equine Influenza Epidemiology Support Group (2008). Equine influenza 2007. The Australian Experience: Report to the Consultative Committee on Emergency Animal Disease. Department of Agriculture Forestry and Fisheries Canberra

Firestone, S.M., Christley, R.M., Ward, M.P. and Dhand, N.K. (2012). Adding the spatial dimension to the social network analysis of an epidemic: Investigation of the 2007 outbreak of equine influenza in Australia. Prev. Vet. Med. 106, 123-135

Firestone, S.M., Lewis, F.I., Schemann, K.A., Ward, M.P., Toribio, J.A.L.M.L. and Dhand, N.K. (provisionally accepted for publication). Understanding the associations between on-farm biosecurity practice and equine influenza infection during the 2007 outbreak in Australia. Prev. Vet. Med.

Firestone, S.M., Schemann, K.A., Toribio, J.-A.L.M.L., Ward, M.P. and Dhand, N.K. (2011). A case-control study of risk factors for equine influenza spread onto horse premises during the 2007 epidemic in Australia. Prev. Vet. Med. 100, 53-63

Foord, A.J., Selleck, P., Colling, A., Klippel, J., Middleton, D. and Heine, H.G. (2009). Real-time RT-PCR for detection of equine influenza and evaluation using samples from horses infected with A/equine/Sydney/2007 (H3N8). Vet. Microbiol. 137, 1-9

Friedman, N., Goldszmidt, M. and Wyner, A. (1999). Data analysis with Bayesian networks: A bootstrap approach. In: Kaufmann, M. (Ed.), Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence (UAI'99), San Francisco, USA, 196-205

Guthrie, A.J., Stevens, K.B. and Bosman, P.P. (1999). The circumstances surrounding the outbreak and spread of equine influenza in South Africa. Rev. Sci. Tech. 18, 179-185

Hand, D.J., McConway, K.J. and Stanghellini, E. (1997). Graphical models of applicants for credit. IMA J Management Math 8, 143-155

Heckerman, D., Geiger, D. and Chickering, D.M. (1995). Learning Bayesian networks - The combination of knowledge and statistical data. Mach. Learn. 20, 197-243

Hernan, M.A., Hernandez-Diaz, S., Werler, M.M. and Mitchell, A.A. (2002). Causal knowledge as a prerequisite for confounding evaluation: An application to birth defects epidemiology. Am. J. Epidemiol. 155, 176-184

Hoeting, J.A., Madigan, D., Raftery, A.E. and Volinsky, C.T. (1999). Bayesian model averaging: A tutorial. Stat. Sci. 14, 382-401

Lewis, F.I. (2012). Additive Bayesian network modelling in R. Available at: http://www.r-bayesian-networks.org/. Accessed 26 November 2012

Lewis, F.I., Brulisauer, F. and Gunn, G.J. (2011). Structure discovery in Bayesian networks: An analytical tool for analysing complex animal health data. Prev. Vet. Med. 100, 109-115

Lewis, F.I. and McCormick, B.J.J. (2012). Revealing the complexity of health determinants in resource-poor settings. Am. J. Epidemiol. 176, 1051-1059

MacKay, D.J.C. (1992). Bayesian interpolation. Neural Comput. 4, 415-447

Maddux, J.E. and Rogers, R.W. (1983). Protection motivation and self-efficacy - A revised theory of fear appeals and attitude change. J. Exp. Soc. Psychol. 19, 469-479

Norman, P., Boer, H. and Seydel, E.R. (2005). Protection motivation theory. In: Conner, M., Norman, P. (Eds.), Predicting Health Behaviour: Research and Practice with Social Cognition Models. 2nd Edition. Open University Press, Maidenhead, 81-126

NSW DPI (2007). Equine influenza protection plan. NSW Dept of Primary Industries. Orange, NSW, Australia

OIE (2012). Manual of diagnostic tests and vaccines for terrestrial animals 2012. Chapter 2.5.7. Equine influenza. Available at: http://www.oie.int/international-standard-setting/terrestrial-manual/. Accessed: 23 November 2012

Poon, A.F.Y., Lewis, F.I., Pond, S.L.K. and Frost, S.D.W. (2007). Evolutionary interactions between N-linked glycosylation sites in the HIV-1 envelope. PLoS Comput. Biol. 3, 110-11

R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, Available at: http://www.R-project.org/ Accessed 20 March 2012

Schemann, K., Firestone, S.M., Taylor, M.R., Toribio, J.A.L.M.L., Ward, M.P. and Dhand, N.K., (2012a). Horse owners'/managers' perceptions about effectiveness of biosecurity measures based on their experiences during the 2007 equine influenza outbreak in Australia. Prev. Vet. Med. 106, 97-107

Schemann, K., Taylor, M.R., Toribio, J. and Dhand, N.K. (2011). Horse owners' biosecurity practices following the first equine influenza outbreak in Australia. Prev. Vet. Med. 102, 304-314

Schemann, K.A., Firestone, S.M., Taylor, M.R., Toribio, J.-A.L.M.L., Ward, M. and Dhand, N.K. (2012b). From the horse's mouth: Perceptions of the management of the 2007 equine influenza outbreak in Australia. Transbound. Emerg. Dis. DOI:10.1111/j.1865-1682.2012.01305.x

Schemann, K.A., Lewis, F.I., Firestone, S.M., Ward, M.P., Toribio, J.-A.L.M.L., Taylor, M.R. and Dhand, N.K. (provisionally accepted for publication). Untangling the complex inter-relationships between horse managers' perceptions of effectiveness of biosecurity practices using Bayesian graphical modelling. Prev. Vet. Med

Taylor, M.R., Agho, K.E., Stevens, G.J. and Raphael, B. (2008). Factors influencing psychological distress during a disease epidemic: Data from Australia's first outbreak of equine influenza. BMC Public Health 8

Taylor, M.R., Stevens, G.J., Agho, K.E., Kable, S.A. and Raphael, B. (2012). Crying wolf? Impact of the H1N1 2009 influenza pandemic on anticipated public response to a future pandemic. Med. J. Aust. 197, 561-554

Tierney, L. and Kadane, J.B. (1986). Accurate approximations for posterior moments and marginal densities. J. Am. Stat. Assoc. 81, 82-86

Umeh, K. (2004). Cognitive appraisals, maladaptive coping, and past behaviour in protection motivation. Psychol. Health 19, 719-735

Uppal, P.K. and Yadav, M.P. (1987). Outbreak of equine influenza in India. Vet. Rec. 121, 569-570

Watson, J., Selleck, P., Axell, A., Bruce, K., Taylor, T., Heine, H., Daniels, P. and Jeggo, M. (2011). Diagnosis of equine influenza virus infections in quarantine stations in Australia, 2007. Aust. Vet. J. 89, 4-6

Society for Veterinary Epidemiology and Preventive Medicine

PAST VENUES AND ORGANISERS OF ANNUAL MEETINGS

Year	**Venue**	**Organiser(s)**
1983	Southampton	Davies & Thrusfield
1984	Edinburgh	Thrusfield
1985	Reading	Thrusfield
1986	Edinburgh	Thrusfield
1987	Solihull	Thrusfield
1988	Edinburgh	Thrusfield
1989	Exeter	Howe
1990	Belfast	McIlroy
1991	London	Jones
1992	Edinburgh	Thrusfield
1993	Exeter	Howe
1994	Belfast	Menzies
1995	Reading	Paterson
1996	Glasgow	Reid
1997	Chester	Clarkson
1998	Ennis, Ireland	Collins
1999	Bristol	Green
2000	Edinburgh	Thrusfield & Mellor
2001	Noordwijkerhout, The Netherlands	van Klink
2002	Cambridge	Wood & Newton
2003	Warwick	Green
2004	Martigny, Switzerland	Stärk
2005	Nairn	Gunn
2006	Exeter	Peeler
2007	Dipoli, Finland	Virtala & Alban
2008	Liverpool	Pinchbeck & Robinson
2009	London	Verheyen & Pfeiffer
2010	Nantes, France	Fourichon & Hoch
2011	Leipzig, Germany	Thulke & Lange
2012	Glasgow	Parkin & Others
2013	Madrid, Spain	Martínez-López & Vargas Alvarez

PAST PRESIDENTS

1982-’83	G. Davies
1983-’84	P.R. Ellis
1984-’85	G. Gettinby
1985-’86	R.W.J. Plenderleith
1986-’87	M.V. Thrusfield
1987-’88	K.S. Howe
1988-’89	A.M. Russell
1989-’90	S.G. McIlroy
1990-’91	J.E.T. Jones
1991-’92	J.M. Booth
1992-’93	E.A. Goodall
1993-’94	R.G. Eddy
1994-’95	J.T. Done
1995-’96	G.J. Gunn
1996-’97	M.S. Richards
1997-’98	J.D. Collins
1998-’99	F.D. Menzies
1999-’00	K.L. Morgan
2000-’01	S.W.J. Reid
2001-’02	A.D. Paterson
2002-’03	L.E. Green
2003-’04	J.L.N. Wood
2004-’05	E.G.M. van Klink
2005-’06	D.J. Mellor
2006-’07	E. J. Peeler
2007-’08	J. R Newton
2008-’09	L. Alban
2009-’10	D.U. Pfeiffer
2010-’11	L.A. Kelly
2011-’12	C. Fourichon

EXECUTIVE COMMITTEE 2012-2013

T. Parkin (President), C. Fourichon (Senior Vice-President), K. Verheyen (Junior Vice-President), K. Mintiens (Honorary Secretary), M. Brennan (Honorary Treasurer), A. Lindberg, L. Rosenbaum Nielsen, H-H Thulke, T. Martinez (co-opted), B. Martínez-López (co-opted) and S. More (co-opted)

Honorary Auditors: Dominic Mellor & Fraser Menzies

LIFE MEMBERS

J.M. Booth, M.J. Clarkson, J.D. Collins (deceased), G. Davies, J.T. Done (deceased), R.G. Eddy, P.R. Ellis, E.A. Goodall, G. Gettinby, K.S. Howe, M.E. Hugh-Jones, W. Martin, F. Menzies, A.M. Russell, M.V. Thrusfield, J. Wilesmith, K. Morgan

PLENARY TALKS

Year	Gareth Davies Lecture	Conference Opening Plenary
2013	Andreas Hensel Dioxins, EHEC and strawberries: Risk assessment and risk communication in practice	José Manuel Sánchez-Vizcaíno The Spanish experience on the control and eradication of infectious diseases: from the old to the current system
2012	Stuart Reid Evidence-based prevention: well done or rare	Didier Boichard Genomic selection: an opportunity for improving health of farm animals
2011	Karin Schwabenbauer From science to policy - the case of classical swine fever (CSF) control	Dominic Mellor The trouble with epidemiology: the tyranny of numbers
2010	David Waltner-Toews Beyond one world, one health and ecohealth…what's out there?	James Wood From pathogen adaption to host ecology: epidemiological and experimental contributions to the understanding of emerging infectious diseases
2009	Jørgen Westergaard The interaction between veterinary science, legislation and management in animal disease control in the European Union	Katharina Stärk Food safety challenges in a global market – are we ready?
2008	Paul Fine Infectious disease eradication – meanings and implications	Kenton Morgan For the benefit of Mr Kite
2007	Yrjö Gröhn Food supply veterinary medicine: Modelling of production, health and food safety	Laura Green Improving Animal Health

2006	David Galligan From partial budgets to real options - concepts in animal health economics	Nigel French Understanding human exposure to zoonoses from food and the environment: The application of molecular tools and modeling
2005	Bill Reilly: From TB to VTEC: The changing epidemiology of foodborne zoonoses	Simon More: Towards eradication of bovine tuberculosis in Ireland: A critical review of progress
2004	Ulrich Kihm: BSE and the stable to table concept	Gary Smith: Spatial models of infectious disease in the USA: a crisis of conference and confidentiality
2003	Sir David Cox: The current state of statistical science	Ynte Schukken: Molecular and mathematical epidemiology of bovine mastitis
2002	George Gettinby: Informatics and epidemiology – the first 400 years	Bryan Grenfell: Deterministic and stochastic influences on the dynamics and control of infectious diseases
2001	Will Houston: Science politics and animal health policy: epidemiology in action	Mart de Jong: Design and analysis of transmission experiments
2000	Jim Scudamore: Surveillance – past, present and future	Dirk Pfeiffer: Spatial analysis – a new challenge for veterinary epidemiologists
1999	Aalt Dijkhuizen: The 1997/98 outbreak of classical swine fever in the Netherlands: lessons learned from an economic perspective	Mark Woolhouse: Understanding the epidemiology of scrapie
1998	Wayne Martin: Art, science and mathematics revisited: the role of epidemiology in promoting animal health	

SOCIETY FOR VETERINARY EPIDEMIOLOGY AND PREVENTIVE MEDICINE

APPLICATION FOR MEMBERSHIP

Name ....

Address ....

....

....

....

Telephone: ....

Fax: ....

E-mail: ....

Signed .... Date ....

Please enclose the membership fee (£40 sterling to cover two years' membership) along with this application form. Overseas members without British bank accounts are requested to pay 2 - 3 years in advance. Cheques should be in £ sterling and drawn from a British bank. British members should pay future dues by standing order (forms are available from the Secretary or Treasurer). Payment can also be made by credit card; the appropriate form is available from the Society's website, http://www.svepm.org.uk/, or from the Secretary or Treasurer.

Please send this form to the Society's Treasurer:

Dr Marnie Brennan
School of Veterinary Medicine and Science
Sutton Bonington Campus
College Road
Sutton Bonington
Leicestershire LE12 5RD
UK

TEL +44 (0) 115 951 6577
FAX +44 (0) 115 951 6415
Email: marnie.brennan@nottingham.ac.uk

Please turn over

INTEREST GROUPS

Please tick appropriate boxes to indicate your interests:

☐ Analytical Epidemiology (Observational Studies)

☐ Quantitative Epidemiology & Statistical Techniques (Incl. Statistical/Mathematical Modelling)

☐ Herd/Flock Level Disease Control Strategies

☐ National/International Disease Control Policy

☐ Sero-Epidemiology

☐ Herd Health and Productivity Systems

☐ Disease Nomenclature and Epidemiological Terminology

☐ Economic Effects of Disease on Animal Production

☐ Veterinary Public Health and Food Hygiene

☐ Computing, including data logging

☐ Computer Programming *per se*

☐ Population and Animal Disease Databases

☐ Information System Design

☐ Geographical Information Systems (GIS)

☐ Risk Analysis

CONSTITUTION AND RULES

NAME

1. The society will be named the Society for Veterinary Epidemiology and Preventive Medicine.

OBJECTS

2. The objects of the Society will be to promote veterinary epidemiology and preventive medicine.

MEMBERSHIP

3. Membership will be open to persons either actively engaged or interested in veterinary epidemiology and preventive medicine.
4. Membership is conditional on the return to the Secretary of a completed application form and subscription equivalent to the rate for one calendar year. Subsequent subscriptions fall due on the first day of May each year.
5. Non-payment of subscription for six months will be interpreted as resignation from the Society.

OFFICERS OF THE SOCIETY

6. The Officers of the Society will be President, Senior Vice-President, Junior Vice-President, Honorary Secretary and Honorary Treasurer. Officers will be elected annually at the Annual General Meeting, with the exception of the President and Senior Vice-President who will assume office. No officer can continue in the same office for longer than six years.

COMMITTEE

7. The Executive Committee of the Society normally will comprise the officers of the Society and not more than four ordinary elected members. However, the Committee will have powers of co-option.

ELECTION

8. The election of office bearers and ordinary committee members will take place at the Annual General Meeting. Ordinary members of the Executive Committee will be elected for a period of three years. Retiring members of the Executive Committee will be eligible for re-election. Members will receive nomination forms with notification of the Annual General Meeting. Completed nomination forms, including the signatures of a proposer, seconder, and the nominee, will be returned to the Secretary at least 21 days before the date of the Annual General Meeting. Unless a nomination is unopposed, election will be by secret ballot at the Annual General Meeting. Only in the event of there being no nomination for any vacant post will the Chairman take nominations at the Annual General Meeting. Tellers will be appointed by unanimous agreement of the Annual General Meeting.

FINANCE

9. An annual subscription will be paid by each member in advance on the first day of May each year. The amount will be decided at the annual general meeting and will be decided by a simple majority vote of members present at the Annual General Meeting.

10. The Honorary Treasurer will receive, for the use of the Society, all monies payable to it and from such monies will pay all sums payable by the Society. He will keep account of all such receipts and payments in a manner directed by the Executive Committee. All monies received by the Society will be paid into such a bank as may be decided by the Executive Committee of the Society and in the name of the Society. All cheques will be signed by either the Honorary Treasurer or the Honorary Secretary.
11. Two auditors will be appointed annually by members at the Annual General Meeting. The audited accounts and balance sheet will be circulated to members with the notice concerning the Annual General Meeting and will be presented to the meeting.

MEETINGS

12. Ordinary general meetings of the Society will be held at such a time as the Executive Committee may decide on the recommendations of members. The Annual General Meeting will be held in conjunction with an ordinary general meeting.

GUESTS

13. Members may invite non-members to ordinary general meetings.

PUBLICATION

14. The proceedings of the meetings of the Society will not be reported either in part or in whole without the written permission of the Executive Committee.
15. The Society may produce publications at the discretion of the Executive Committee.

GENERAL

16. All meetings will be convened by notice at least 21 days before the meeting.
17. The President will preside at all general and executive meetings or, in his absence, the Senior Vice-President or, in his absence, the Junior Vice-President or, in his absence, the Honorary Secretary or, in his absence, the Honorary Treasurer. Failing any of these, the members present will elect one of their number to preside as Chairman.
18. The conduct of all business transacted will be under the control of the Chairman, to whom all remarks must be addressed and whose ruling on a point of order, or on the admissibility of an explanation, will be final and will not be open to discussion at the meeting at which it is delivered. However, this rule will not preclude any member from raising any question upon the ruling of the chair by notice of motion.
19. In case of an equal division of votes, the Chairman of the meeting will have a second and casting vote.
20. All members on election will be supplied with a copy of this constitution.
21. No alteration will be made to these rules except by a two-thirds majority of those members voting at an annual general meeting of the Society, and then only if notice of intention to alter the constitution concerned will have appeared in the notice convening the meeting. A quorum will constitute twenty per cent of members.
22. Any matter not provided for in this constitution will be dealt with at the discretion of the Executive Committee.

April, 1982
Revised March, 1985; April, 1988; November 1994
Corrected January 1997; April 2002