SOCIETY FOR VETERINARY EPIDEMIOLOGY AND PREVENTIVE MEDICINE

Proceedings of a meeting held at

Liverpool, UK

on the 26th – 28th March 2008

Edited by E.J. Peeler, L. Alban, A. Russell and the SVEPM Executive Committee

ISBN 978-0-948073-84-7

ACKNOWLEDGEMENTS

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

National Centre for Zoonosis Research

Department for Environment, Food & Rural Affairs

Petplan Charitable Trust

Faculty of Veterinary Science, University of Liverpool

Veterinary Laboratories Agency

The Royal Veterinary College

CONTENTS

BIOSECURITY

EPIDEMIOLOGICAL METHODS

SURVEILLANCE

FOOT AND MOUTH DISEASE

***

Society for Veterinary Epidemiology and Preventive Medicine

EQUINE EPIDEMIOLOGY

CASE-CONTROL STUDY TO INVESTIGATE RISK FACTORS FOR IMPACTION COLIC IN DONKEYS IN THE UK

R. COX*, F. BURDEN, L. GOSDEN, C. PROUDMAN, A. TRAWFORD AND G. PINCHBECK

SUMMARY

Impaction colic is the single most common type of colic diagnosed in a large population of donkeys (more than 2000 animals) at The Donkey Sanctuary, UK and the fatality rate from the disease is high. This paper reports risk factors for impaction colic in this population during 2006, identified using an unmatched case-control study. There were 71 cases of impaction colic and multivariable analysis identified a number of variables associated with the disease. Management factors that increased the risk of impaction included paper bedding, feeding of concentrates, limited access to pasture and increasing number of carers. In addition, health variables that were associated with an increased risk of impaction colic were weight loss, recent vaccination and a number of dental abnormalities. The implications of these associations are described and possible practical measures that might be implemented to reduce the risk of the disease are discussed.

INTRODUCTION

Previous research has identified that colic commonly affects the health and welfare of donkeys at The Donkey Sanctuary, and that impaction colic (due to impacted ingesta in the large intestine) is one of the most common types of colic diagnosed (Duffield et al., 2002a, b; Cox et al., 2007). Retrospective analysis of The Donkey Sanctuary clinical database has shown that impaction colic was implicated in more than half of the colic episodes seen in the population of more than 4500 donkeys that were Sanctuary residents between January 2000 and March 2005 and more than half did not recover from the disease (Cox et al., 2007).

Colic is a serious problem in other equid populations, having a detrimental effect on morbidity and welfare (Archer & Proudman, 2006). Impaction colic can be the most commonly occurring type of colic in some populations (Mair & Hillyer, 1997; Brosnahan & Paradis, 2003b). Research has demonstrated that the causes are complex and there are conflicting views about the impact of individual risk factors. Some risk factors are well documented e.g. change in diet (Tinker et al., 1997; Hillyer et al., 2002), while other suggested associations with impaction colic, e.g. dental disease (Hillyer et al., 2002), require further investigation. In addition, previous work has demonstrated that some types of colic, including impaction colic, are seasonal in occurrence (Archer et al., 2006) and further work is required to elucidate potential reasons for this. Although several risk factors for impaction colic have been identified in horses, differences

*Ruth Cox, Faculty of Veterinary Science, University of Liverpool, Leahurst, Neston, Wirral, CH 64 7TE. Email: r.cox@liverpool.ac.uk

in management, physiology and behaviour between the donkey and horse may mean these cannot be extrapolated to donkeys.

Only recently has epidemiological research been conducted to identify specific risk factors for impaction colic in donkeys. Work published elsewhere identified a number of risk factors for impaction colic, including increasing age, feeding extra concentrates and previous episodes of colic (Cox et al., 2007). These results concur with previous studies of colic in other equids. Farm was also identified as a significant risk factor for impaction colic. The study also suggested that dental disease predisposed donkeys to intestinal obstruction and that there was a seasonal effect on the occurrence of impaction colic. Due to the retrospective nature of this study, a number of other factors were not considered due to missing data (e.g. less than half of the donkeys had had a dental exam in the 6 months before the case of colic) or because the information was not routinely recorded in the existing database.

The aim of this study was to identify and quantify risk factors for impaction colic in UK donkeys using a prospective case-control study of all animals housed at The Donkey Sanctuary during 2006. This will aid with identification of high-risk individuals and may also highlight areas where intervention strategies may be introduced to reduce the incidence of the disease.

MATERIALS AND METHODS

Study design

A prospective unmatched case-control study was conducted to identify risk factors for impaction colic in donkeys housed at The Donkey Sanctuary from 1st January 2006 until 31st December 2006. Sample size estimation performed using Win Episcope 2.0 showed that, with 4 controls per case, a study with 65 cases would have 80% power to detect odds ratios of 2.5 or more with 95% confidence, for exposures of 20% in the control population. Previous reports (Duffield et al., 2002a; Cox et al., 2007) showed that at least 65 cases of impaction colic should be obtained during a 12-month period. To ensure that cases and controls were unmatched on time, 20 controls were recruited during each month of the study.

Identification of cases and control animals

The study population consisted of all donkeys housed at The Donkey Sanctuary farms, Devon, UK. Cases were all donkeys that suffered impaction colic diagnosed by a veterinary surgeon at The Donkey Sanctuary, either by rectal examination or at post mortem examination, during the study period. Controls were selected randomly from the population throughout the study. At the start of each month, the number of donkeys present at the Sanctuary was calculated. Each donkey was ordered according to its Animal ID number (assigned when the donkey arrived at the Sanctuary) and control animals were selected using random number generation (Microsoft Excel). Controls were excluded if they had suffered colic in the 14 days prior to selection.

Data collection

A specific questionnaire form, for completion by Donkey Sanctuary personnel, was designed with input from Donkey Sanctuary staff before the trial began and a pilot study was performed.

The questionnaire requested detailed information about all cases and controls (Table 1). One section (2 pages, 12 questions) was completed by the veterinary surgeon when a case of impaction colic was diagnosed. This section requested information about diagnosis, treatment and outcome of all cases. The second section of the questionnaire (10 pages) was completed for all cases and controls by one of the authors (FB) at the Sanctuary and requested information about donkey management and health. In addition, all cases and controls received a detailed dental examination. A total of 90% of the dental examinations were conducted by the resident dental technician (LG); the remainder were conducted by the rsident pathologist and by 1 veterinary surgeon at The Donkey Sanctuary. Examinations were conducted as close as possible to being selected as a case or control (within a maximum of 7 weeks) and this information was recorded on a dental record chart.

Data analysis

The number of cases of impaction colic that occurred during the study period were summed and the incidence rate was calculated as the number of new episodes of colic per 100 donkey years at risk. Multiple episodes were included in this analysis since the aim was to estimate the incidence rate of colic events; a new episode was recorded when the donkey had been free from colic in the previous 14 days. Denominator data were obtained by summing the number of donkey years at risk in the 12-month period.

Information about the 71 cases of impaction colic was compared to information about the 247 control animals. We collected information about management practices in the 2 weeks before the colic case and in the time 2 to 4 weeks before the colic case. Information about the previous 2 weeks was used in analysis, since the management of donkeys did not tend to change between 4 and 2 weeks and because the 2 weeks prior to colic provided information about the time closest to the illness.

Screening of all variables was performed using univariable logistic regression to assess the effects of all variables on the outcome of impaction colic. Categorical variables with small numbers of observations in 1 or more categories, or where the reference category contained relatively few individuals, were recoded to create fewer categories or to create a different reference category. Prior to inclusion in the final multivariable logistic regression model, the functional form (shape) of the relationship between continuous variables and the outcome of colic were explored using generalised additive models (GAM) and smoothing splines (Hastie & Tibshirani, 1990). The functional forms of the continuous variables were then used to inform the fit of these variables in the multivariable logistic regression model. Piece-wise terms, categorised variables and polynomial terms were all considered until the most parsimonious model was achieved.

Variables with $P<0.25$ in the univariable analysis were considered for inclusion in multivariable logistic regression models. Due to the large number of variables eligible for inclusion in multivariable modelling ($P<0.25$), several sub-models were initially created: donkey information, daily routine and housing, feeding, change in routine, health and preventive medicine and dentition. Variables were retained in each model if they significantly improved the fit of the model (assessed by the likelihood ratio test statistic $P<0.05$). The variables identified in each of these models were pooled and used to develop the final effects model (Reeves 1996). As it was not possible to fit all data into 1 multivariable model a stepwise approach was used with the criteria for inclusion of a variable being a significant influence on the fit of the model as assessed by the change in deviance. All the variables initially considered for inclusion ($P<0.25$)

were then forced back into the model to ensure that no significant or confounding variables had been excluded in the model building process. To allow for clustering within location, farm ID was included in the multivariable model as a random effect.

Table 1. Information collected about each case and control in a case-control study of impaction colic in donkeys housed at The Donkey Sanctuary, UK, in 2006

VARIABLE	DESCRIPTION
Donkey details	Age, gender, weight, height, heart girth, body condition score, years spent at Sanctuary, farm location (n=8 farms)
Management	Type of housing, type of bedding, type of outside area, number of carers, number of other animals in enclosure.
Daily routine	Number of hours spent inside and outside, access to pasture and hedgerows, access to playpen.
Behaviour	Occurrence and timing of stereotypies.
Feeding	Type of forage, type and amount of concentrate feed, any other type of extra feed, supplements, number of times fed per day, method of feeding, time since extra rations began, consumption of non-food items, number and type of water source.
Exercise and transport	Frequency of specific activities (e.g. riding), frequency of transport.
Change in routine	Change in: stabling or routine, time at pasture, companion animals, exercise, transport, health or treatment, feeding, behaviour.
Health and preventive health care	Any medical problems, routine examinations and treatment, mobility score, eye sight score, parasite count, parasite pasture management.
Dental disease	Occurrence and position of: missing teeth, loose teeth, hooks, calculus, displaced, broken, sharp, worn, abscess, ulceration, shear, wave, step, undershot, overshot, diastema.
Impaction colic diagnosis (by veterinarian)	Date of colic, method of diagnosis, identification of site of impaction, treatment, presenting signs, outcome, any other clinical abnormalities.

Logistic regression was performed using maximum likelihood estimates conducted in EGRET (Egret for Windows 2.0, Cytel Software Corporation, 1999). All other analyses were conducted in SPSS version 12 (SPSS Inc 2003).

RESULTS

Cases of impaction colic

There were 71 cases of impaction colic in 69 donkeys, in the population of approximately 1800 donkeys housed at the Sanctuary during 2006. The incidence rate was 3.9 impaction colics

per 100 donkeys per year. The majority of impactions (81.7%, n=58) were diagnosed by rectal examination; 16 of these were diagnosed in combination with post mortem findings or clinical signs. The remainder (18.3%, n=13) were diagnosed at post mortem only. The site of impaction was diagnosed in all but 1 case; the majority of impactions (51%) occurred in the pelvic flexure (Table 2).

A total of 48% of donkeys (n=34) recovered from their impaction colic following treatment; none recovered without veterinary treatment. The remaining 52% were euthanased; 81.1% (n=30) of these were euthanased due to the colic episode, while the remainder (18.9%) were euthanased due to a combination of factors.

A number of presenting signs were recorded when donkeys were suffering from impaction colic, most commonly they were described as 'dull' and showed reduced appetite or anorexia. A total of 66.2% of the donkeys (n=47) were reported to have at least 1 other clinical abnormality before the colic episode.

Table 2. Site of impaction in 71 cases of impaction colic in donkeys at The Donkey Sanctuary, UK, in 2006

IMPACTION SITE	NUMBER OF CASES
Pelvic flexure	41
Caecum	14
Small colon/rectum	11
Large colon	7
Other	3
Unidentified	1
Total	77[a]

[a]Total number of cases = 71, however in 6 cases, impaction occurred concurrently in 2 different sites (pelvic flexure + large colon in 2 cases; pelvic flexure + small colon/rectum in 1 case; pelvic flexure + caecum in 2 cases; large colon + small colon in 1 case).

Univariable analysis

Univariable analysis identified a large number of variables that significantly ($P<0.05$) influenced the occurrence of impaction colic. Body condition score, weight and heart girth all showed that lighter or lower condition animals were at increased risk. Feeding of extra concentrate rations was associated with an increase in risk and many individual concentrate feed types were associated with this increase e.g. chaff, alfalfa, high fibre cubes and cereals. The feeding of straw was associated with decreased risk. Donkeys without water sources in their outside enclosure were at increased risk of colic. Donkeys without water access in the inside enclosure were also at increased risk but this was not significant, possibly due to the very small number of donkeys that did not have water available inside. Housing also influenced risk of colic, with animals housed in a single stable at increased risk and conversely those in a barn at decreased risk.

Many of the variables indicating that the donkey had experienced a change in routine in the previous 4 weeks were associated with increased risk. These included a change in health status, change in behaviour and change in both the amount and type of concentrate feed.

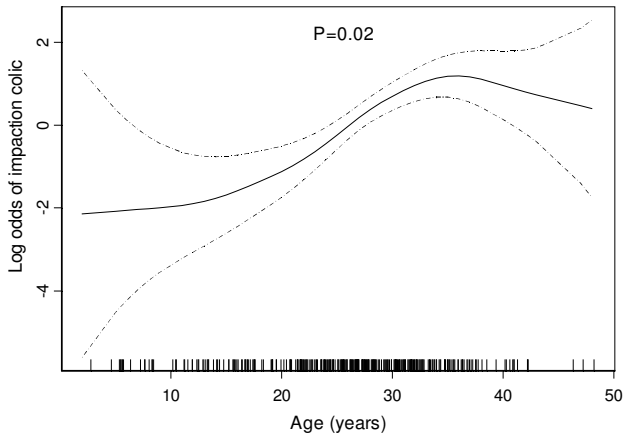
A large proportion of animals (63% of cases and 25% of controls) had suffered from a health problem in the 4 weeks prior to having colic or being selected as a control. Colic, weight loss, poor eye sight, laminitis, lameness or hoof problems and appearing lethargic or dull or exhibiting other behavioural changes in the previous 4 weeks were all associated with increased risk of impaction.

Dental disease was also very common; 92.1 % (293 donkeys) had at least 1 dental disease or abnormality. The most common abnormalities were missing, loose or displaced teeth, dental hooks, sharp teeth and diastemata. Many of these were associated with an increased risk of impaction in univariable analysis.

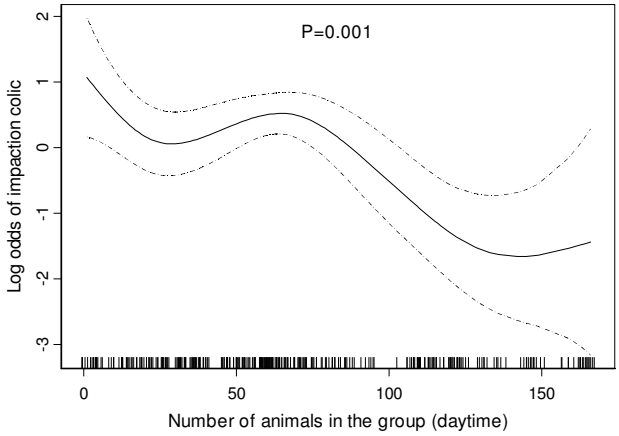
Continuous variables that were significant in univariable analysis included age of the donkey, the number of animals that shared the enclosure and the number of missing teeth (both incisors and cheek teeth). Results from the GAMs are shown in Fig. 1. There was a linear increase in risk with increasing age up until approximately 38 years and then the risk levelled off (Fig. 1a). The greater the number of animals in the shared enclosure the lower the risk. There was a linear decrease in risk with increasing numbers of animals but there was an increase around approximately 60-70 animals (Fig. 1b). There was a linear increase in risk with increasing numbers of missing cheek teeth up to 13 or 14 teeth, after this the risk levelled off (Fig. 1c) and this variable was fitted in the multivariable model as a piece-wise linear term with this form. There were very few donkeys with more than 13 missing cheek teeth.

Multivariable analysis

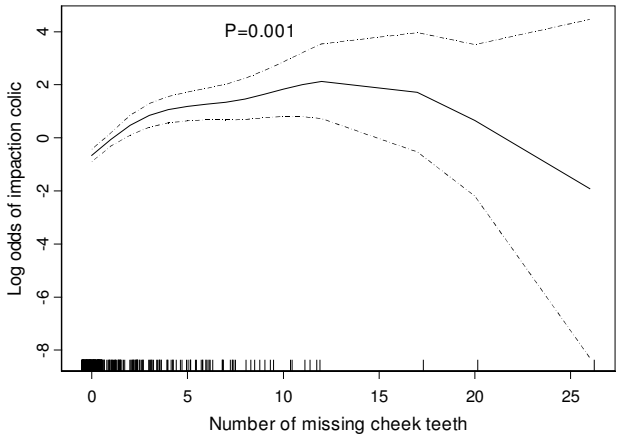
The final multivariable model demonstrated that a number of donkey and management variables were associated with the risk of impaction colic in this population (Table 3). This model included farm as a random effect, however the farm-level variation was zero after the inclusion of fixed effects suggesting that these fixed effects accounted for any clustering within farms.



a. Age



b. Number of animals in same enclosure during the day



c. Number of missing cheek teeth

Fig. 1 Functional forms of the continuous variables modelled in univariable generalised additive models (continuous fixed effects fitted using smoothers) to determine relationship between the predictor variable and the outcome (log odds of colic). Plots show the fitted curves with 95% confidence intervals (dashed lines). Number of data points represented by x-axis rug plot. Chi-square P-value for non-linearity shown for significantly non-linear variables.

Table 3. Final multivariable model of risk factors associated with impaction colic in donkeys at The Donkey Sanctuary, UK, in 2006

VARIABLE		ODDS RATIO	LOWER 95% CI[a]	UPPER 95% CI[a]	LRS[b] P-VALUE
Number of carers	1-2	1			
	3-4	4.44	1.1	17.9	
	5-6	11.14	2.5	49.5	
	7-10	4.76	1.1	20.9	0.01
Paper bedding	No	1			
	Yes	12.56	2.5	64.3	0.002
Concentrate rations	No	1			
	Yes	5.17	1.6	16.4	0.005
Weight loss	No	1			
	Yes	25.68	3.1	210.4	0.003
Vaccination	No	1			
	Yes	12.71	1.7	97.1	0.01
Pasture access	24 hours	1			
	Some	2.39	0.7	8.1	
	None	3.44	1.3	8.8	0.04
Hooks on teeth	No	1			
	Yes	0.15	0.03	0.8	0.03
Worn teeth	No	1			
	Yes	3.84	1.5	10.0	0.006
Ulcers	No	1			
	Yes	16.36	1.7	159.5	0.02
Diastemata	0	1			
	1	1.66	0.5	5.2	
	2	2.57	0.9	7.7	
	3+	9.97	3.2	30.6	<0.001
Missing cheek teeth[c]		1.16	1.0	1.3	0.009

[a]CI = Confidence Interval; [b]LRS = Likelihood ratio test statistic; [c]Fitted as a piece-wise linear term with linear increase in risk up until 14

The greater the number of carers responsible for the donkey, the greater the increase in risk; donkeys with 1 or 2 carers were at lowest risk. Donkeys that were fed extra rations as concentrate feed were at a five-fold increased risk of impaction. In addition, animals with no access to pasture were at increased risk (odds ratio 3.4) compared to those with 24-hour access. Donkeys bedded on paper were at increased risk of colic, however this had wide confidence intervals due to the small number of cases and controls that were bedded on paper. These animals were in the hospital and at 2 other farms. Animals that were vaccinated in the previous 2 weeks were also at increased risk of colic. Animals that experienced weight loss in the previous 4 weeks were at increased risk of colic although few animals had this health problem. Other variables relating to body condition did not remain in the multivariable model. A number of dental pathologies were significantly associated with impaction colic in the multivariable model;

these included missing cheek teeth, ulcers, diastemata and worn teeth which all increased the risk of colic. The presence of hooks was still associated with decreased risk of colic.

After allowing for the above variables, age was no longer significant (P=0.9) and forcing age into the model did not change the effects of the other variables. Forcing month (or season) into the final model showed that this was not significant, demonstrating that this model explained some of the seasonality.

DISCUSSION

This study identified all veterinary-diagnosed cases of impaction colic in a population of more than 1800 donkeys in the UK over a 12 month period. The incidence rate of impaction colic in this population was similar to previous retrospective work at The Donkey Sanctuary which found an incidence of 3.2 episodes per 100 donkeys per year over a five-year period (Cox et al., 2007). The case fatality rate was almost identical to the 51% fatality rate recorded during retrospective analysis of cases in the same study. This fatality rate is high compared to other equid populations, where, for example, 95% of horses treated medically and 58% of horses treated surgically survived for 1 year after large colon impaction (Dabareiner & White, 1995). The case fatality rate could be an overestimate if some donkeys experienced a mild impaction and recovered without diagnosis or treatment, and because the estimate includes cases that were only confirmed at post mortem; however, it is still much greater than other reports. Differences may be explained by the old age of the donkeys, that the decision for euthanasia considered the health of other body systems and that other equids may be treated quite differently; for example, surgery is performed more routinely in horses (Reeves et al., 1989; Proudman, 1991; Brosnahan & Paradis, 2003b). In addition, diagnosis of colic in donkeys can be more difficult than in horses since accurate and localised detection of pain in donkeys is difficult because of the subtlety of presenting signs (Taylor & Matthews, 1998; Ashley et al., 2005), and abdominal pain may not be diagnosed until the later stages of the disease (Duffield et al., 2002a).

This study identified a number of variables associated with impaction colic. Donkeys with 1 or 2 carers were at decreased risk. This may be due to more dedicated care in feeding or management, with possible increased identification of colic in donkeys in small groups with few carers. Similarly, horses whose owners provide their care are at decreased risk of colic (Reeves et al., 1996; Hillyer et al., 2001) and this has also been attributed to more dedicated care.

In this study, as in a retrospective study at The Donkey Sanctuary (Cox et al., 2007), the provision of concentrate rations was a significant risk factor for impaction colic. Univariable analysis demonstrated that many of the commonly used concentrate feeds seemed to be associated with increased risk of impaction and none appeared protective. Feed type has been identified as a cause of increased risk of colic in some studies of equines (Hudson et al., 2001) although others have reported no association with type of concentrate (Cohen et al., 1999; Traub-Dargatz et al., 2001). The amount of concentrate (Tinker et al., 1997; Hudson et al., 2001) has been identified as a significant risk factor for colic, with the highest amount being associated with the greatest risk. Furthermore, any recent change in the type or amount of concentrate can be associated with increased risk (Tinker et al., 1997; Cohen et al., 1999; Hudson et al., 2001; Hillyer et al., 2002). Univariable analysis indicated that both a recent change in amount and type of concentrate feed was associated with increased risk of impaction although these variables did not remain in the final multivariable model.

Donkeys with limited or no access to pasture were at increased risk of impaction. Limited pasture access has also been identified as a risk factor for colic in other equine populations and this could be the result in feeding changes that interfere with normal intestinal motility (Hudson et al., 2001). Horses fed less easily digested, more complex or varied diets with a high proportion of forage in the form of hay or pasture have been reported to be at decreased risk of colic (Tinker et al., 1997). It has been suggested that more time at pasture may decrease the risk by allowing for continuous grazing, therefore avoiding the physiological problems induced by feeding concentrate meals twice daily (Clarke et al., 1990). Furthermore, the feeding of, or bedding on, straw has been associated with impaction in horses. However, in this study very few donkeys were not fed straw and there was no evidence that feeding straw increased the risk of colic.

Recent changes in access to pasture can also influence the occurrence of impaction, and univariable analysis showed that a change in pasture was significantly associated with impaction colic. This concurs with other studies that have found that a recent decrease in pasture availability (either no time at pasture, or a decrease in pasture acreage or time at pasture) was a risk factor for colic (Hudson et al., 2001).

The association between access to pasture and colic may also be an indication of activity since colic has been associated with activity level and changes in activity levels (Cohen et al., 1995; Kaneene et al., 1997; Cohen et al., 1999). Adequate exercise can be important in normal large intestine function, with exercise causing an increase in feed digestibility in horses (Orton et al., 1985) and in donkeys (although this result was not statistically significant) (Pearson and Merritt, 1991). Horses that spent an increased amount of time in a stall can be at increased risk of large colon impaction (Dabareiner & White, 1995) and horses that are exercised at least once per week can be at a greater risk of colic relative to horses turned out to pasture only (Cohen et al., 1999). Results in the present study indicate that one possible way to reduce the risk of colic in those donkeys needing additional nutrition, and in high risk animals, is the provision of 24-hour access to pasture.

Reduced access to water in enclosures was associated with impaction in univariable analysis; this has been suggested to be a risk factor for colic in general (Reeves et al., 1996; White, 1997) and specifically impaction colic in horses (Pugh and Thompson, 1992). Although this variable did not remain in the final model, this may be due to the small numbers of animals without access to water in both enclosures. This is a relatively simple preventive measure to implement.

Paper or cardboard bedding was used in the Sanctuary hospital and on two farms. On one of these farms where the bedding was made from cardboard there were reports of donkeys eating this bedding and it has been found in the stomach of donkeys at post mortem (F. Burden, personal communication). Animals may be at risk of impaction due to being hospitalised, as has been shown in horses (Senior et al., 2006). In addition, hospitalisation is likely to indicate a concurrent health problem in the donkey. Indeed, donkeys housed on paper or cardboard bedding at the Sanctuary are often those with respiratory problems.

It is not clear why vaccination in the previous 2 weeks increased the risk of colic, although this may be due to the process of vaccination affecting the donkeys' routine – e.g. they may be separated from companions or confined. Many of the variables associated with a change in routine were associated with colic in univariable analysis.

This study confirmed previous evidence that dental pathology is associated with impaction colic in donkeys (Cox et al., 2007). A number of dental pathologies that increased the risk of impaction were identified, namely diastemata, missing cheek teeth, ulcers and worn teeth. Previous studies have suggested a link between dental disease and impaction colic in donkeys (Duffield et al., 2002a, b) and colic in geriatric equidae (Brosnahan & Paradis, 2003a, b), although none of these studies identified particular dental abnormalities in detail. Other studies on younger horses have also noted an association between large colon impaction or obstruction and poor dentition (White, 1997) or infrequent dental treatment (Hillyer et al., 2002).

While it is difficult to understand why the presence of hooks appeared to reduce the risk of impaction colic (odds ratio 0.8), it is of note that this was the only abnormality that was significantly more common in younger animals, while the other abnormalities were more common in the older animals. It was also noted that hooks are seen more often in animals that recently arrived at the Sanctuary compared to those that have been resident for some time (F. Burden, personal communication). We hypothesise therefore that hooks may be a marker for relatively normal dentition in this population. It is possible that hooks may not be severe enough to cause any kind of eating or chewing problem to the extent of influencing digestibility of feed.

Many of the dental disorders described can only be prevented, rather than cured, and implementation of quality preventive dental care as early as possible in the donkey's life may be beneficial. The Donkey Sanctuary has already instigated this at their farms by the recruitment of 2 dedicated dental technicians. In addition, education regarding donkey dentition to a wider group (e.g. veterinary surgeons and dental technicians) may be useful and has also been implemented by the Sanctuary.

Having an existing health problem and having mobility or vision problems were all associated with colic in univariable analysis. Only dental disease and weight loss remained in the final model. This may be due to the small numbers of animals that suffered from each specific medical problem in the 4 weeks prior to entry to the study. However, this finding does provide some further evidence that donkeys suffering from other diseases are at increased risk of impaction. Although many of these may not be preventable they do help to identify animals at high risk of impaction.

After inclusion of the fixed effects variables in the final model neither age, nor month or season were significant. This suggests that some of the age and seasonality effects were explained by the variables measured in this study. For example, dental pathology and health status (animals suffering from weight loss) may be better explanatory variables for identifying donkeys at high risk than simply using age alone. This may be one reason why, in general, there are conflicting views about the association between age and colic (Archer & Proudman, 2006). Number of carers, bedding type and pasture access may all explain some of the seasonality demonstrated in the retrospective study by Cox et al. (2007).

This study has identified variables which may help to identify donkeys at high risk of impaction colic; those with a history of weight loss and with concurrent dental pathology. These animals may need particular care in terms of management. In addition, early identification of impaction colic, which may be difficult in donkeys due to the subtlety of the presenting signs, and hence early treatment, may help improve the survival rate of donkeys with impactions. The study also identified some variables that may be targeted to reduce the incidence of impaction colic in this donkey population, for example reduced concentrate feeding and access to pasture. Further work is required to test these preventive measures, on farm, in controlled trials.

ACKNOWLEDGEMENTS

We acknowledge the vets and staff of The Donkey Sanctuary for help with data collection. This study was funded by The Donkey Sanctuary.

REFERENCES

Archer, D.C., Pinchbeck, G.L., Proudman, C.J. and Clough, H.E. (2006). Is equine colic seasonal? Novel application of a model based approached. BMC. Vet. Res. 2:27, 1-33

Archer, D.C. and Proudman, C.J. (2006). Epidemiological clues to preventing colic. Vet. J. 172, 29-39

Ashley, F.H., Waterman-Pearson, A.E. and Whay, H.R. (2005). Behavioural assessment of pain in horses and donkeys: application to clinical practice and future studies. Equine Vet. J. 37, 565-575

Brosnahan, M.M. and Paradis, M.R. (2003a). Assessment of clinical characteristics, management practices, and activities of geriatric horses. J. Am. Vet. Med. Assoc. 223, 99-103

Brosnahan, M.M. and Paradis, M.R. (2003b). Demographic and clinical characteristics of geriatric horses: 467 cases (1989-1999). J. Am. Vet. Med. Assoc. 223, 93-98

Clarke, L.L., Roberts, M.L. and Argenzio, R.A. (1990). Feeding and digestive problems in horses, physiological responses to a concentrated meal. Vet. Clin. North. Am. Equine Pract. 6, 433-449

Cohen, N.D., Gibbs, P.G. and Woods, A.M. (1999). Dietary and other management factors associated with colic in horses. J. Am. Vet. Med. Assoc. 215, 53-60

Cohen, N.D., Matejka, P.L., Honnas, C.M. and Hooper, R.N. (1995). Case-control study of the association between various management factors and development of colic in horses. J. Am. Vet. Med. Assoc. 206, 667-673

Cox, R., Proudman, C.J., Trawford, A.F., Burden, F. and Pinchbeck, G.L. (2007). Epidemiology of impaction colic in donkeys in the UK. BMC. Vet. Res. 3

Dabareiner, R.M. and White, N.A. (1995). Large colon impaction in horses: 147 cases (1985-1991). J. Am. Vet. Med. Assoc. 206, 679-685

Duffield, H.F., Bell, N. and Henson, F.M.D. (2002a). The common presenting signs and causes of colic in the donkey. In: BEVA (Ed.), Proceedings of the 41st British Equine Veterinary Association Congress, Glasgow, UK, p. 214

Duffield, H.F., Bell, N. and Henson, F.M.D. (2002b). Factors associated with impactive colic in the donkey. In: BEVA (Ed.), Proceedings of the 7th International Equine Research Symposium, Manchester, UK, p. 122

Hastie, T. and Tibshirani, R. (1990). Generalized additive models. Chapman and Hall London

Hillyer, M.H., Taylor, F.G.R. and French, N.P. (2001). A cross-sectional study of colic in horses on thoroughbred training premises in the British Isles in 1997. Equine Vet. J. 33, 380-385

Hillyer, M.H., Taylor, F.G.R., Proudman, C.J., Edwards, G.B., Smith, J.E. and French, N.P. (2002). Case control study to identify risk factors for simple colonic obstruction and distension colic in horses. Equine Vet. J. 34, 455-463

Hudson, J.M., Cohen, N.D., Gibbs, P.G. and Thompson, J.A. (2001). Feeding practices associated with colic in horses. J. Am. Vet. Med. Assoc. 219, 1419-1425

Kaneene, J.B., Miller, R., Ross, W.A., Gallagher, K., Marteniuk, J. and Rook, J. (1997). Risk factors for colic in the Michigan (USA) equine population. Prev. Vet. Med. 30, 23-36

Mair, T.S. and Hillyer, M.H. (1997). Chronic colic in the mature horse: A retrospective review of 106 cases. Equine Vet. J. 29, 415-420

Orton, R.K., Hume, I.D. and Leng, R.A. (1985). Effects of exercise and level of dietary protein on digestive function in horses. Equine Vet. J. 17, 386-390

Pearson, R.A. and Merritt, J.B. (1991). Intake digestion and gastro-intestinal transit time in resting donkeys and ponies and exercised donkeys given *ad libitum* hay and straw diets. Equine Vet. J. 23, 339-343

Proudman, C.J. (1991). A two year, prospective survey of equine colic in general practice. Equine Vet. J. 24, 90-93

Pugh, D.G. and Thompson, J.T. (1992). Impaction colics attributed to decreased water intake and feeding coastal bermuda grass hay in a boarding stable. Equine Pract. 14, 9-12

Reeves, M.J., Gay, J.M., Hilbert, B.J. and Morris, R.S. (1989). Association of age, sex breed factors in acute equine colic: A retrospective study of 320 cases admitted to a veterinary teaching hospital in the USA. Prev. Vet. Med. 7, 149-160

Reeves, M.J., Salman, M.D. and Smith, G. (1996). Risk factors for equine acute abdominal disease (colic): results from a multi-centre case-control study. Prev. Vet. Med. 26, 285-301

Senior, J.M., Pinchbeck, G.L., Allister, R., Dugdale, A.H.A., Clark, L., Clutton, R.E., Coumbe, K., Dyson, S. and Clegg, P.D. (2006). Post anaesthetic colic in horses: a preventable complication? Equine Vet. J. 38, 479-484

Taylor, T.S. and Matthews, N.S. (1998). Mammoth asses - selected behavioural considerations for the veterinarian. Appl. Anim. Beh. Sci. 60, 283-289

Tinker, M.K., White, N.A., Lessard, P., Thatcher, C.D., Pelzer, K.D., Davis, B. and Carmel, D.K. (1997). Prospective study of equine colic risk factors. Equine Vet. J. 29, 454-458

Traub-Dargatz, J.L., Kopral, C.A., Seitzinger, A.H., Garber, L.P., Forde, K. and White, N.A. (2001). Estimate of the national incidence of and operation-level risk factors for colic among horses in the United States, spring 1998 to spring 1999. J. Am. Vet. Med. Assoc. 219, 67-71

White, J.L.N. (1997). Risk factors associated with colic. In: Robinson, N. (Ed.), Current Therapy in Equine Medicine. W. B. Saunders Co, Philadelphia, pp. 174-179

COST OF VETERINARY CARE IN INSURED SWEDISH HORSES 1997-2004

A. EGENVALL*, B.N. BONNETT, S. LARSDOTTER AND U. EMANUELSON

SUMMARY

Costs registered for veterinary care in non-racing horses, with complete insurance at a Swedish insurance company covering >30% of the Swedish horse population, were studied over time (1997-2004). Associations with gender, age, breed group, geographical location and life-insurance value were studied. The univariable focus was on yearly costs per case and per horse-year at risk (HYAR). Tobit regression was used to assess the yearly relationship between the log-transformed costs and breed, gender, age, geographical location and life-insurance value, censoring the data at the deductible level. Poisson regression was used analogously, but with number of reimbursed cases as the outcome and including the log of HYAR as offset. The total population was 141,552 horses contributing a total time at risk of 498,119 HYAR, on average 62,265 per year. In total ~37,000 horses had insurance claims. The incidence rates in the years 1997, 2000 and 2004 were 1227, 1282 and 1080 claimed horses per 10,000 HYAR. Costs per claimed horse increased from 4905 SEK in 1997 to 7805 SEK in 2004, compared to costs per HYAR of 571 SEK and 805 SEK for the same years, respectively. The difference in SEK per HYAR between sub-categories was largest for life-insurance value (<15,000 and ≥45,000 SEK; 377 and 1652 SEK/HYAR respectively) and least for gender (geldings and stallions; 807 and 520 SEK/HYAR). The estimates from the Tobit and Poisson regressions showed, in general, similar patterns, except for the youngest horses where the Tobit models showed low estimates and the Poisson high relative risk ratios. Even though neither of the models behaved very well at validation, it is believed that they aided in simultaneous evaluations of the effects. In the present data, comparing figures from 1997 and 2004, the increase in costs per claimed horse was 59% and the increase in cost per HYAR was 41%, compared to a consumer price index increase of 9.8%.

INTRODUCTION

There is limited information on the cost of illness of riding horses and even direct veterinary costs have not been well described. A study on the national incidence of equine colic in the US in 1998 found that the mean cost of colic was US$160 per event, while for surgical cases it was US$3872. The mean number of days lost per colic event was 2.2 days and of the estimated total cost (US$115.3 million), death loss accounted for 66% (Traub-Dargatz et al., 2001). Cohen et al. (2004) found an almost double cost for cases with postoperative ileus compared to those without, using a material of 251 surgically treated horses. Abutarbush and Naylor (2005) further found that the cost of treatment of nephrosplenic entrapment of the large colon in 19 horses was

*Agneta Egenvall, Department of Clinical Sciences, Faculty of Veterinary Medicine and Animal Husbandry, Swedish University of Agricultural Sciences, SE-750 07 Uppsala, Sweden. Email: agneta.egenvall@kv.slu.se

significantly higher in surgically than medically treated horses (~ 5 times higher). Furthermore, Mason et al. (2005) compared costs of "standing non-sutured" to "recumbent sutured" castration and found that the cost of the first alternative was one third of the latter, irrespective of a higher complication rate of the former. The overall median of cost for colic per horse was US$4.42 in a prospective one-year study in the early nineties (Lloyd and Kaneene, 1997). However, costs of illness in horses include the direct costs for investigations and treatment delivered by veterinarians and other health professionals (i.e. blacksmiths, non-veterinarian alternative medicine practitioners). Substantial costs further accrue from days lost for work as well as replacement costs, if horses have to be put down or cannot be used fully for the intended purpose. In racing horses, earnings continue to be determined in relation to certain disease conditions (e.g. Schnabel et al., 2007), however the treatment costs are seldom included.

It is possible to study the direct costs of veterinary care (investigations and treatments delivered by a veterinarian) using claims data from insurance companies for events where costs exceeded the deductible. For example, using data from a Swedish animal insurance database (www.agria.se), average yearly costs per insured horse for reimbursed veterinary care for locomotor problems during a 5-year follow-up period varied from 880 to 1132 SEK in a cohort with locomotor-problem claims in 1997 and between 410 and 580 SEK in a cohort without locomotor claims in 1997 (Egenvall et al., 2008).

The objective of this study was to describe gross direct costs registered for veterinary care in all non-racing horses with complete insurance at Agria over time (1997-2004), taking into account effects of gender, age, breed, geographical location and life-insurance value.

MATERIALS AND METHODS

The insurance database

Data for this study came from the insurance company Agria that covers approximately 30% of the horse population in Sweden (Egenvall et al., 2005). A complete insurance for veterinary care reimburses the owner if the horse becomes wounded or ill for various reasons. Costs reimbursed include examination and treatment delivered by a veterinarian as well as medications used at the veterinary visit. Complete insurance does not cover travel costs of the veterinarian or costs for prescribed medications. In 2000, the yearly premium per horse for a complete insurance for veterinary care varied between 275 SEK and 950 SEK (on average, from 1997-2004, one SEK was ~ 0.117 US$). The yearly fee for veterinary care insurance is not related to the value of the horse but instead depends on which premium group the horse belongs to (for further details see Egenvall et al., 2005). The deductible for a claim period (90-100 days) was approximately 1098 SEK, with the owner paying 20% of veterinary costs that exceeded the deductible. The maximum sum of reimbursement in 1997 was 20,000 SEK; in 2001 this increased to 25,000 SEK and in 2004 to 50,000 SEK, above which expenses were fully paid by the owner. Using the insurance does not influence coverage or fees. When the veterinary care insurance is settled, the information on principal diagnosis, as well as the net and gross costs from the receipts for veterinary care are entered into the insurance database by a clerk processing the claim. With life insurance, the owner generally gets reimbursed if the horse dies or is euthanased. Almost all insured horses have both veterinary care and life insurance. Note that actively racing horses cannot have complete insurance coverage. In general, one visit at a veterinary clinic corresponds to one receipt and one entry in the database.

Data management

Data were extracted from the insurance company's computerised database and details on data management have been reported previously (Egenvall et al., 2005; Egenvall et al., 2006a). Data on gender, date of birth, breed, postal code of owner, identification of the horse, information on whether the horse was covered for veterinary care and/or life, value for which the horses were life-insured ("life insurance value"), dates of death/euthanasia, dates of visits to veterinarians and costs for all veterinary-care event claims were used. Within a year, a horse was said to be "claimed" when any receipts for veterinary care had been submitted and registered in the database. Accordingly, the cost was the total (gross) from submitted receipts, regardless of the amount that was reimbursed. Dates when horses entered or left the insurance program and reasons why they were removed were also available.

Horses were assigned the age they had on the 1st of January each year. Horses were defined to have lived in southern, middle, or northern Sweden according to the owner's postal code. Horses were also defined as "urban" (horses whose owners had a postal code in or close to one of the three major cities) or "rural", thus identifying whether horses lived in an area of higher or lower human population density. The three and two levels of region and urban, respectively, were combined. However, for the combination of northern-urban there were no horses because there are no large cities in northern Sweden. Life insurance value was categorised into three levels: ≤10,000 SEK, >10,000 to ≤30,000 SEK, >30,000 to ≤45,000 SEK and >45,000 SEK.

Study population

The study population comprised all non-racing horses with complete insurance for veterinary care any time between 1997 and 2004 at Agria with one exception: horses >22 years of age at the beginning of each year were excluded.

Descriptive statistics

Descriptive calculations were done for each year. The univariable focus for costs was on cost per horse-year at risk (HYAR) and cost per claimed horse. A horse was denoted as claimed when it had ≥1 veterinary-care receipt submitted within a year and all costs on all claims within that year for a specific horse made up the cost. Incidence rate (IR) calculations were carried out with the exact time at risk as the denominator. Also, incidences were calculated for each year. The yearly descriptive statistics and incidences were also averaged over the eight years and denoted as "average" figures.

Tobit and Poisson regression

Tobit regression (QLIM procedure in SAS software, SAS Institute Inc., Cary, NC, 27513, USA) was used to assess the relationship between the cost (divided by 1000 and then log-transformed) and breed, gender, age, geographical location and life-insurance value within year. Baselines were categories where the averaged costs per HYAR were lowest. Median-centered second-order polynomials were used for the HYAR per horse and life-insurance value. Age was included as a categorical variable, based on inspection of the average age-specific IRs, with the following categories; -1<0 years (horses born during the year), 0<1, 1<2, 2<3, 3<4, 4<15 and >15 years of age (baseline). In the Tobit models, data were considered censored at the deductible (from 1100 SEK in 1997 to 2100 SEK in 2004). P-values <0.05 were regarded as statistically significant and were evaluated using the T-test.

Poisson regression (GENMOD procedure in SAS) was used to assess the relationship between the number of reimbursed cases (horses with a claim) and the same variables as used in the Tobit analysis. However, in the Poisson, HYAR was not included as a fixed factor; the logarithm of HYAR was instead used as an offset. P-values were evaluated using the type 3 criterion.

To be able to compare between models the focus was on main effects models. For the Poisson model the ratio of deviance to the degrees of freedom has been reported. Model validation was done using residual Q-Q plots and inspection of outliers (data from 1997 and 2004). With respect to life insurance value, sensitivity analysis was performed including observations with a life insurance value <200,000 as well as <400,000 SEK on data from 1997 and 2004.

RESULTS

Study population

The total population was 141,552 horses contributing a total of 498,119 HYAR, varying yearly between 59,396 HYAR (in 2001) and 67,339 HYAR (in 2004). The median number of HYAR that each horse contributed was 2.7, varying from <0.1 to 8. During the study period, 37,247 horses had insurance claims. Table 1 shows the distribution, by HYAR, for the categorical variables. The mean of the median and mean ages were 8 and 8.4 years, respectively.

Incidence rates and costs

Table 1 demonstrates the averaged IRs, costs per HYAR and costs per claimed horse by gender, breed group, geography and life insurance value. The incidence rate varied from 1227 claimed horses per 10,000 HYAR in 1997 to 1080 claimed horses per 10,000 HYAR in 2004. Costs per claimed horse increased from 4905 SEK in 1997 to 7805 SEK in 2004 (Fig. 1), compared to costs per HYAR of 571 SEK and 805 SEK for the same years, respectively. Figure 2 shows age-specific averaged IRs, costs per HYAR and costs per claimed horse, while the same by year and life insurance value is given in Fig. 3. On average 4% of the total costs were from rejected receipts.

Multivariable results

Table 2 shows the averaged Tobit and Poisson regression models. In the Tobit regressions, all covariates were significant; at least one p-value from the T-test was <0.05. The Q-Q plots showed both curvature and lack of small values in the middle. However, large proportions of the residuals were exactly zero (≥90%). Looking at these, none had reimbursed costs over the deductible, otherwise the distribution of the variables (gender, age, breed group, geography, HYAR) were similar to the total dataset. The life insurance value was just slightly lower in this group. The high residuals were all for individuals with high gross costs, often over 40,000 SEK. The individuals with the lowest residuals all had high life insurance values and most were warmbloods; a proportional number of those, compared to the whole dataset, were reimbursed above the deductible. It was noted that there were few low residuals (below zero) in the years 2002-2004 compared to 1997-2001.

In the Poisson models all p-values for included covariates were <0.0001; however, life insurance value squared was thrown out of two models (no model was estimated when this was

included and life insurance value was completely removed). The ratio of the deviance to the degrees of freedom varied between 0.48 and 0.54, indicating under-dispersion. The Q-Q plot of the deviance residuals showed linear departure in the middle. The largest positive residuals were for warmblood cases with very high life insurance values and the largest negative residuals were for non-cases with high life insurance values. The deviation in the middle of the Q-Q plots (residuals between -0.5 to 1 in 2004) included very few cases and relatively fewer warmbloods, but e.g. age and life insurance values were average.

Table 1. Incidence rates (IR, all claims) and average yearly costs, per claimed horse and per 10,000 horse-years at risk (HYAR) for the analysed variables. Variables are presented with average standard errors (SE). The population is horses insured for veterinary care at Agria[a] during 1997-2004

	IR	SE	HYAR[a]	Cost per HYAR	Cost per claimed horse	SE
Total	1179	14	62265	692	6177	75
Gender						
Stallions	806	32	8536	520	6754	258
Females	1094	19	30214	641	6178	113
Geldings	1429	25	23515	807	6052	108
Breed group						
Not defined	1016	86	1446	551	5677	493
Icelandic horses	590	35	5170	307	5393	325
Coldbloods	477	35	3899	249	5291	559
Ponies	775	19	21738	397	5281	127
Standardbreds	866	52	3259	467	5621	372
Thoroughbreds	1448	62	3978	839	6244	272
Warmblood	1859	30	22775	1140	6712	106
Geography						
Middle-rural	1147	24	21028	640	5859	128
Middle-urban	1496	55	5180	926	6670	246
North-rural	1289	61	3720	590	4823	204
South-rural	1588	53	5958	952	6376	210
South-urban	1036	20	26039	641	6500	129
Life-insurance value						
<15,000 SEK	675	15	29396	377	5677	147
≥15-30,000 SEK	1317	27	19022	744	6057	124
≥30-45,000 SEK	1956	48	9489	1101	6242	147
≥45,000 SEK	2660	85	4358	1652	7144	214

[a] The number of HYAR reported are for the whole period that the horses stayed insured

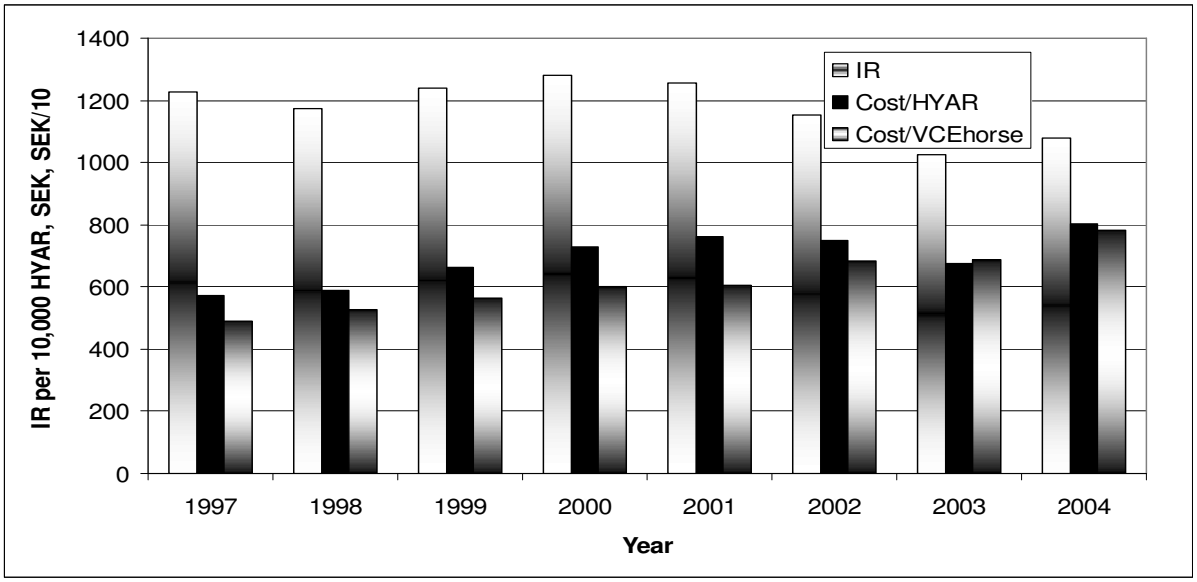


Fig. 1 Yearly incidence rates (IR), costs by horse-years at risk (HYAR) and costs per claimed horse

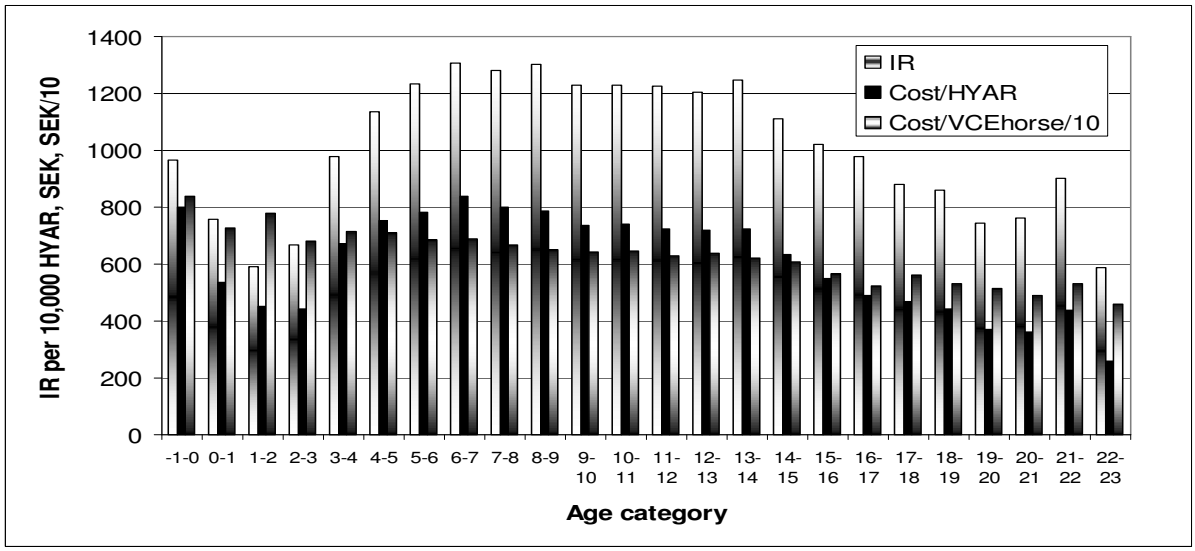


Fig. 2 Incidence rates (IR), costs by horse-years at risk (HYAR) and costs per claimed horse, by age

Comparing estimates from the Tobit and Poisson regressions (exponentiated to relative risk, RR), these in general followed each other well. For example, diagrams for age (Fig. 4) showed a similar pattern for costs (Tobit) and risk of being a claimed horse (Poisson), except however for the youngest horses. For breed group and geography the pattern was similar (Fig. 5). Figure 6 demonstrates the life insurance value estimates, using a few selected life insurance values, and RRs from the models, varying rather widely between the years (two of the Tobit estimates are highly truncated).

Table 2. Mean estimates and standard errors (SE) from the year-stratified multivariable Tobit and Poisson models. Between 68,022 and 76,389 horses were included in the models. The population is horses insured for veterinary care at Agria during 1997-2004

Parameter	Category	Tobit[a]		Poisson[b]	
		Mean estimate	Mean SE	Mean estimate	Mean SE
Intercept		-4.51	0.14	-3.32	0.10
Gender	Mare	0.22	0.06	0.24	0.04
	Gelding	0.44	0.06	0.42	0.05
	Stallion (BL)[c]	0	-	0	-
Age (years at start of year)	-1-0	-0.95	0.13	0.14	0.10
	0-1	-0.26	0.10	-0.06	0.08
	1-2	-0.47	0.11	-0.33	0.09
	2-3	-0.40	0.10	-0.25	0.08
	3-4	-0.02	0.09	0.07	0.07
	4-15	0.26	0.05	0.20	0.03
	>15 (BL)	0	-	0	-
Breed group	Not defined	0.83	0.14	0.69	0.11
	Icelandic horses	0.02	0.11	0.05	0.10
	Ponies	0.47	0.09	0.42	0.08
	Standardbreds	0.51	0.12	0.49	0.10
	Thoroughbreds	1.21	0.11	0.96	0.09
	Warmbloods	1.54	0.09	1.18	0.08
	Coldbloods (BL)	0	-	0	-
Geography	Middle-rural	-0.01	0.07	-0.10	0.05
	Middle-urban	0.15	0.08	-0.02	0.06
	South-urban	0.29	0.08	0.11	0.06
	South-rural	-0.08	0.07	-0.18	0.05
	North-rural (BL)	0	-	0	-
Value[d,e]		0.14	0.01	0.10	0.01
Value squared[f]		-0.002	<0.001	-0.001	<0.001
HYAR[f]		-4.20	0.25	-	-
HYARSQ[g]		-6.00	0.34	-	-

[a] The number of horses with costs varied between 6338 and 7429
[b] The number of un-censored horses varied between 5583 and 6888
[c] BL- baseline
[d] Life-insurance value
[e] Life-insurance value squared
[f] Horse-years at risk
[g] HYAR squared

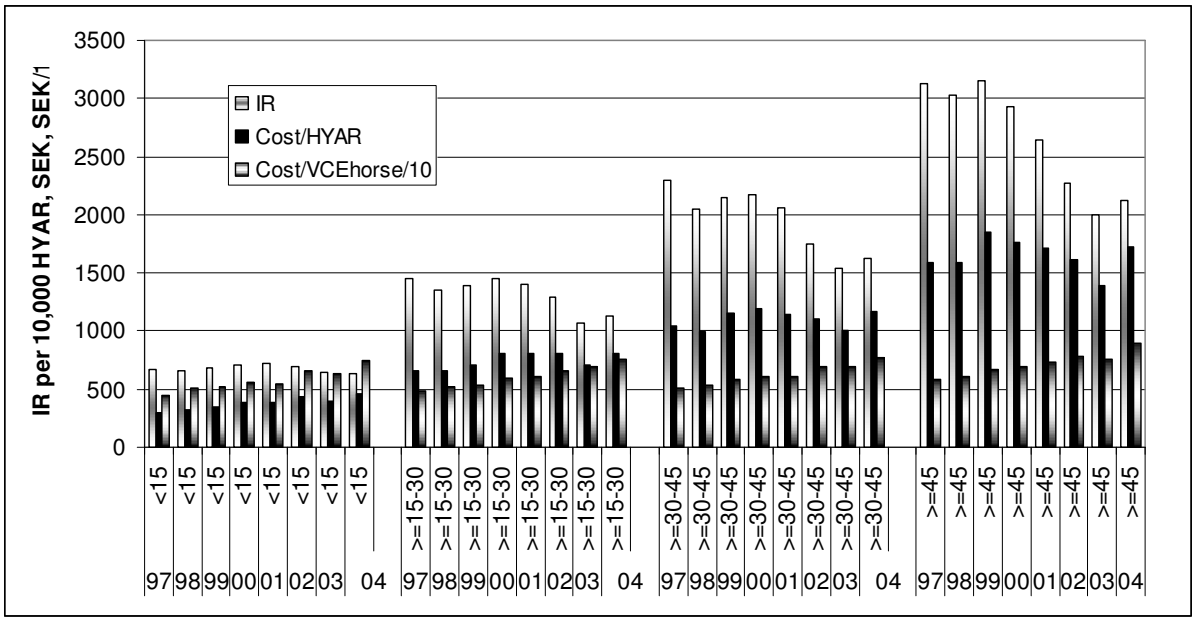


Fig. 3 Incidence rates (IR), costs by horse-years at risk (HYAR) and costs per claimed horse, by year and by life-insurance value

Sensitivity analysis on the 1997 and 2004 data omitted horses with life insurance values ≥200,000 SEK (between 107 and 257 horses each year) and ≥ 400,000 SEK (between 39 and 77 horses each year). Subjective inspection of the results from 1997 in the Tobit model, except for changes related to life insurance value that naturally were seen in all new models, showed mainly deviations for Icelandic horses and horses aged 3-4 and 4-15 years, with substantial decreases of estimates in all cases. In the Poisson model with data from 1997, the estimates for gelding decreased in the new models and the other effects, Icelandic horses and age effects, were similar to those in the Tobit models. In 2004 data, the only major deviance was for Icelandic horses; in both regressions these estimates decreased similarly to those for the 1997 data. A possible reason for the change in the Icelandic horses is that these had the highest life insurance values of all breed groups.

DISCUSSION

Comments on modelling strategies

In the present study the actual cost of veterinary care was the focus, studied using Tobit regression and compared to Poisson regression (Egenvall et al., 2005). Censored Tobit regression is a technique relatively frequently used in econometric modelling, originally developed by Tobin (1958). A typical application is analysis of costs registered when part of the sample has no costs registered or when costs above a threshold are not adequately registered. However, the actual unknown costs may be below or above these thresholds. In the veterinary field it seems that Tobit regression has only been used a few times (Ekstrand and Carpenter, 1998; Chi et al., 2002), apart from in some experimental research.

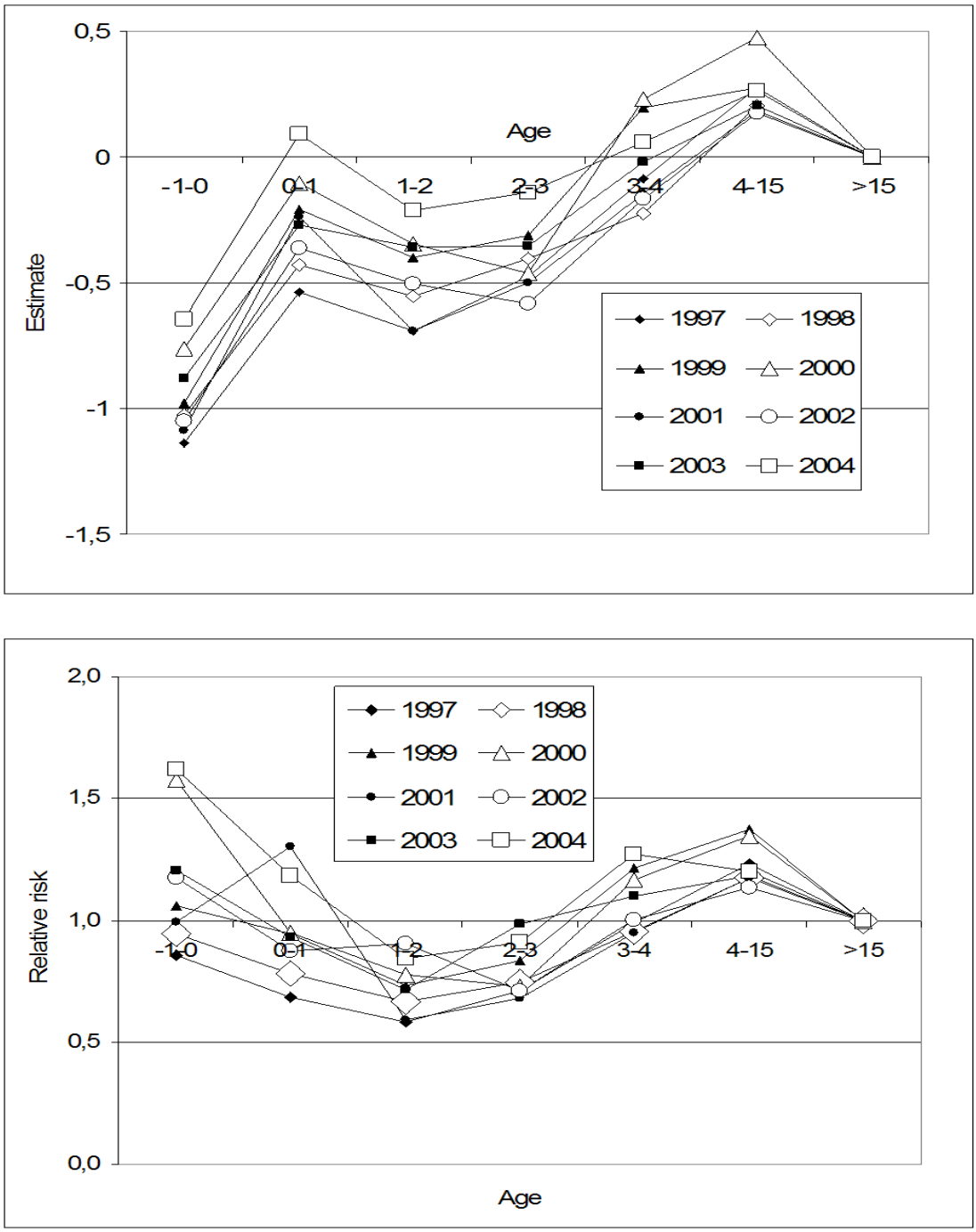


Fig. 4 Year-specific age estimates from Tobit (above) and Poisson (below) regression models. The estimates are controlled for horse-years at risk (HYAR), HYAR squared, life insurance value, life insurance value squared, gender, breed group and geography

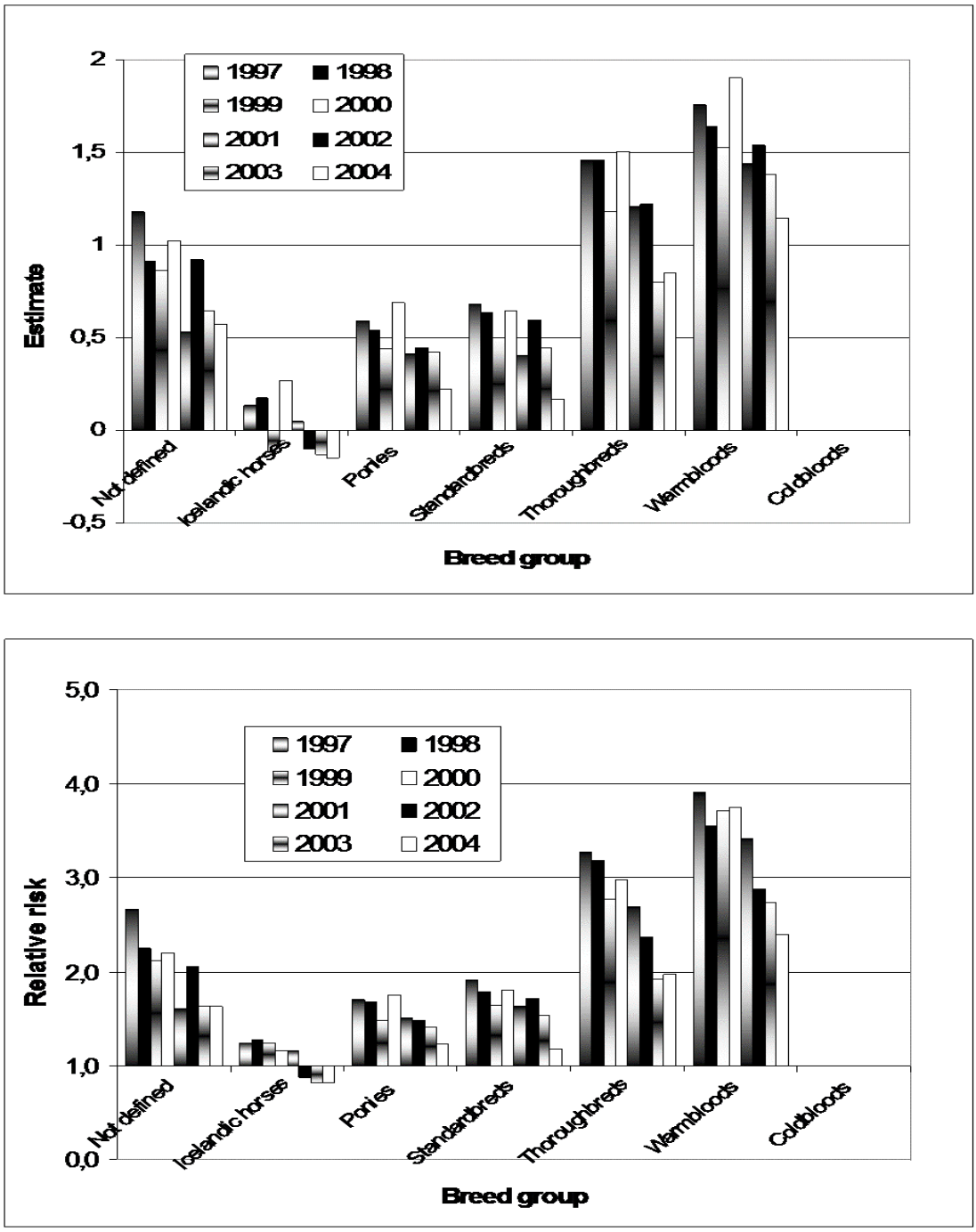
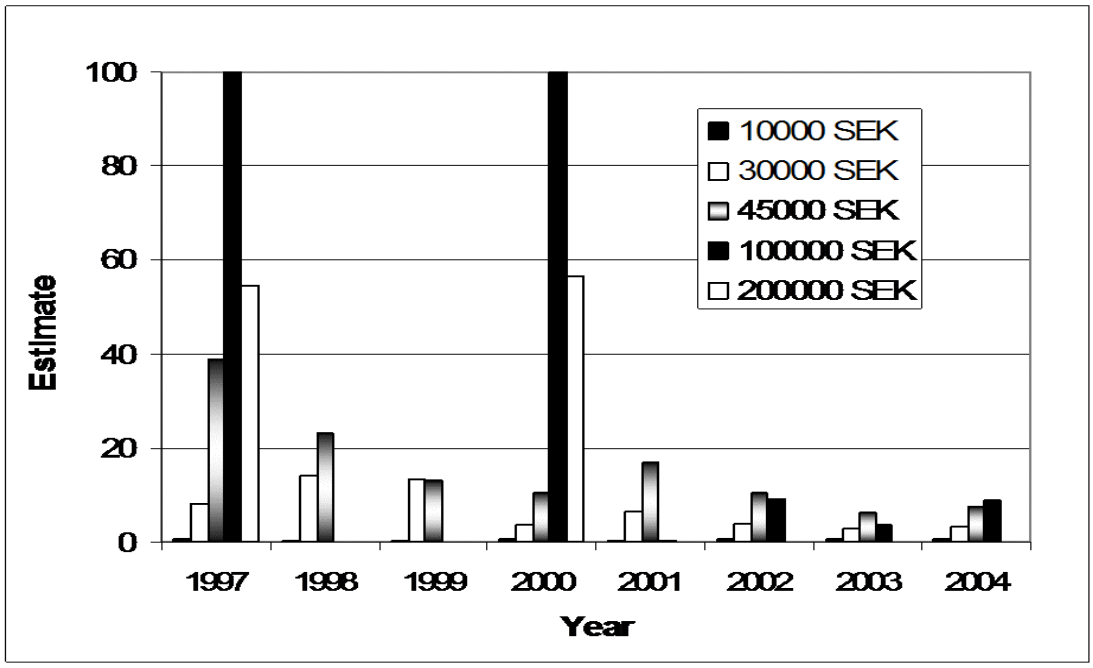


Fig. 5 Year-specific breed group estimates from Tobit (above) and Poisson (below) regression models. The estimates are controlled for horse-years at risk (HYAR), HYAR squared, life-insurance value, life-insurance value squared, gender, age and geography

Tobit regression is estimated using maximum likelihood. This likelihood includes two parts; one part is based on the normal probability density function for observations of uncensored values and the other part, for the censored values, is based on the normal cumulative distribution function (Tobin, 1958; Ekstrand and Carpenter, 1998; Kennedy, 2003).



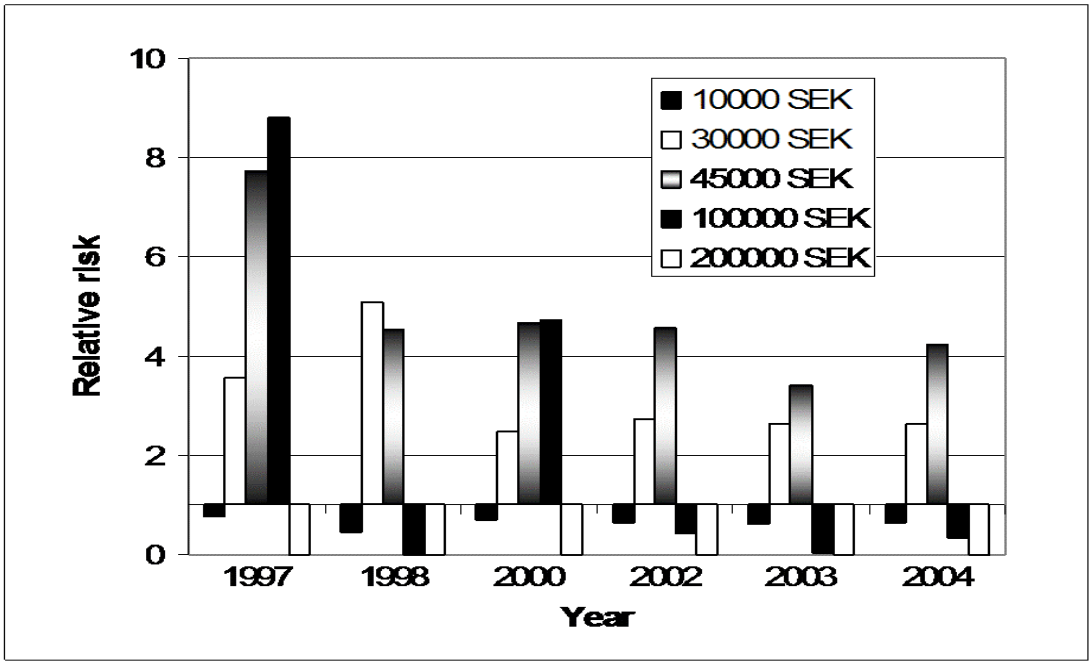


Fig. 6 Year-specific life insurance value estimates from Tobit (above) and Poisson (below) regression models for selected life insurance values. The estimates are controlled for horse-years at risk (HYAR), HYAR squared, geography, gender, age and breed group. Note that estimates in 1997 and 2000 for the 100,000 SEK category are truncated (being 1248 and 138 respectively)

With respect to Tobit regression, it was decided to censor costs at the level of deductible to avoid setting the censored variable much lower than the lower recorded costs (Kennedy, 2003). This led to the stratified data and the Poisson modelling, compared to the Tobit regressions, operating with slightly different "cases" because the cases (claimed horses with costs above the

deductible) were somewhat fewer in the Tobit regressions. Furthermore, a main assumption of Tobit regression is that the process that decides whether a horse becomes claimed is the same as the one that decides the level of the cost (Kennedy, 2003). It appears that the similar results obtained from both multivariable analyses support this concept.

Log-transforming the outcome variable for the Tobit regressions made the Q-Q-plots somewhat better compared to modelling the untransformed variables. Working with predictions from the Tobit models (data not shown), it was found that the model for 2004 had a predictive ability that was much closer to the observed costs if all observations with life insurance values above 60,000 SEK were omitted (omitting 4% of the observations).

The Poisson models were under-dispersed, the reason for which is likely to be lack of data for important determinants. The Q-Q plots were not totally appropriate either, which is likely to be due to the same reason. However, the Poisson and Tobit models agreed well, with a few exceptions (explored below), although life insurance value was sub-optimally modelled (discussed below). The subjective sensitivity analysis omitting horses with large life insurance values was relatively reassuring because estimated effects were relatively stable.

Comparing Tobit and Poisson models between years, the Q-Q plots showed similarities between the years, although there were also deviations. For five years the Q-Q plots of the Poisson models were fairly straight, if the large deviating middle part was ignored (data not shown). For the other three years, including the two years when life insurance value was not included, the Q-Q plots were judged as more bi-linear. In the three later years the Q-Q plots for the Tobit regression almost lacked the left part, i.e. with values below zero, although the right part was still curved (bent to the right; data not shown).

The focus for the analysis of the modelling was on main effect models with the same effects each year. The reason was to provide models that could easily be compared between years, while still controlling for some of the more important factors. Many interactions were significant both at p-value limits of 0.05 and 0.0001 (using data from 1997, 2000 and 2004; data not shown). However, the interactions were not consistent over years; both those that were significant and which categories within the interaction terms had statistical loadings varied between years.

Limitations of the database

It is important to be aware that the costs are not the total veterinary costs delivered to each horse, though we have modelled the “gross costs”. The maximum sum of reimbursement increased by over 100% (from 20,000 SEK to 50,000 SEK) during the time of study. Still, it is likely that a number of disease events exceed the maximum sum of reimbursement, e.g. surgical colic. In such cases it is unlikely that the database will contain full documentation on the veterinary costs accrued. Another problem is that the deductible increased much more than the general costs reflected by the consumer price index. However, in 1997-2000 the deductible and maximum sum of reimbursement remained the same. From Fig. 1 a general upward trend was found during these years, supporting the overall conclusion from the study that costs increased during the study period. It is therefore likely that censoring the data at the deductible attenuated the time effect when comparing Tobit models between years. It is also important to note that the research database not fully reflects the “true” database, because of difficulties in following individual horses (Egenvall et al., 2005).

Time

Figure 1 indicates that the incidence rate went down during the later part of the study period. However, costs by HYAR increased, except in 2003, and cost per claimed horse generally increased. During this time, the Swedish consumer index increased by 9.8% (www.scb.se, accessed 19 December 2007). In the present data, comparing figures from 1997 and 2004, the increase in costs was substantially higher: 59% per claimed horse and 41% per HYAR (data not shown). Obviously, the increased costs are due to increases per claimed horse and not due to more horses being claimed. An explanation is that veterinarians diagnose more diseases in a given patient and pursue treatments to a larger extent in horses that are being investigated because of disease problems and have reached the deductible.

Life-insurance value

Horses with high life-insurance values get relatively more claims and to some extent relatively more expensive claims. The former was reflected by the IR and costs per HYAR and the latter by an increased cost per claimed horse. Life insurance value seems, both from this and previous studies, to be the best proxy for use (e.g. competition) that is found in the insurance database (Egenvall et al., 2006b). It is probable that horses that are used heavily get more injuries, but also that they will be treated for a larger proportion of minor problems (e.g. locomotor problems) because they perform towards the limit of their capacity. Furthermore, a non-competing horse may not have an injury registered by the owner, if only given light exercise, or owners of non-competing horses may be much more willing to give the horse a rest period. Owners of competing horses are likely to be eager to get their horses back to training quickly.

There are some potential problems with the life insurance values as analysed in this study. Each animal had only one life-insurance value used (the first of the different ones found during the period), though some animals will have had their life insurance values updated during the study period (possibly given a new soundness examination and/or no claims for a certain period of time depending on actual increase in value). In the multivariable analysis life insurance value was modelled using median-centered second-order polynomials, but this may not have been adequate because results varied widely between years and overall conclusions were therefore difficult to draw. However, a comparison between models still yielded certain similarities, e.g. the presented values were highest in the same groups and in 1997.

Age, gender, breed and geography

It was found that IRs and costs per HYAR were high in the youngest foals and then decreased to subsequently increase again, as we have previously seen for IRs for veterinary care (Egenvall et al., 2005). Comparing the IRs and costs per HYAR, the costs per HYAR were relatively higher than the IRs, compared to the other age categories. However, the costs per case showed only a slight decline with age, albeit being relatively high in the newborn. The latter was expected, because the disease pattern of young foals makes them likely to receive more aggressive, and thus more costly, medical care in case of illness. Furthermore, horses of an age at which they are most actively used were subject to veterinary care to a relatively large degree; this was also in line with expectations.

The impressions from the multivariable methods were similar except for in the lowest age group. For this age group, the stratified results of IRs and ratios of SEK to HYAR agreed quite

well with the Poisson results, but the Tobit results were less expected. In an attempt to elucidate this, each variable except age was removed from the Tobit model (year 2004), one at a time, and the estimates for this age category compared (see Fig. 4). Removing HYAR, the estimate was at baseline (estimate -0.05 (SE 0.09; p=0.56)). Removing only HYAR squared, the estimate was more negative, but more similar compared to other "low" young age categories (estimate -0.14 (SE 0.09; p=0.05)). In the full model these were -0.65 (SE 0.10; p<0.0001) and in the univariable model -0.28 (SE 0.09; p=0.002). Removing the other variables had no significant effect on the estimate of the lowest age category. Using the same procedure on the 1997 data yielded somewhat less pronounced changes in the same direction (data not shown). It is likely that the modelled HYAR did not fit the youngest age group very well (the estimates of HYAR were widely different if analysed only within the youngest age group, reversed from low negative to high positive). In conclusion, the inclusion of HYAR, and especially HYAR squared, had a pronounced effect on the estimates in this age group. Still, after these manipulations the Tobit estimates never showed the high positive values that the Poisson regressions did.

The gender and breed group effects were significant both in Tobit and Poisson modelling, also in accordance with earlier results (Egenvall et al., 2005). The same gender (gelding) and breed groups (warmbloods and thoroughbreds) were indicated as having the highest risk in both models and all years (part of the data not shown). From Fig. 5 it can be seen that, controlling for other included variables, both the costs and the relative risks within warmbloods and thoroughbreds decreased over time. With respect to the multivariable results for geography, south-urban and to some degree middle-urban had the highest risks (data not shown). This is biologically plausible as horse owners near cities are likely to be using their horses rather heavily (e.g. competition). This is at least partly because it is more expensive to have a horse in urban areas and horse owners have to be rather "motivated for usage" to have a horse stabled in such an area.

Time at risk

HYAR was included as an offset in the Poisson models, hence no specific results were attached to this variable. However, for the Tobit regressions it was decided a priori to force this variable into the models. The estimates were negative and of large magnitude. With HYAR squared added this suggested that costs were unlikely for small HYAR, but more likely for large HYAR, as expected. The inclusion also had substantial effects on some of the age estimates. Both the estimates of the youngest (see example above) and the 3- to 4-year-olds differed if HYAR was included or not. In both cases this was logical because individuals in these categories often entered during the year and contributed little time at risk. An additional complicating factor is the possibility that owners were more eager to use the insurance and find out the problems with their newly bought animal. For other owners the situation might be the reverse: insured animals may be relatively healthier in the period shortly after insurance was taken out, both because they had a recent health check and because the owners were likely to be somewhat less inclined to use the insurance shortly after acquisition (because they want to have a clean start on their insurance period).

CONCLUSION

In general, the estimates from the Tobit and Poisson regressions showed similar patterns, except in the youngest horses where the Tobit models showed low estimates and the Poisson

models high relative risk ratios. Although neither of the models behaved very well at validation, it is believed that they aided in simultaneous evaluations of the effects, leading to better understanding of how factors act together on whether, and the amount of, veterinary-care costs that develop. Costs increased mainly with time and life insurance value; the latter is generally highly related to intensive usage and it was the costs per claimed horse that increased most dramatically.

ACKNOWLEDGEMENTS

The study was funded by grants from the Swedish Foundation for Equine Research. The Agria insurance company is also acknowledged for allowing access to their database.

REFERENCES

Abutarbush, S.M. and Naylor, J.M. (2005). Comparison of surgical versus medical treatment of nephrosplenic entrapment of the large colon in horses: 19 cases (1992-2002). J. Am. Vet. Med. Ass. 227, 603-605

Chi, J., VanLeewen, JA., Weersink, A. and Keefe, G.P. (2002). Management factors related to seroprevalences to bovine viral-diarrhoea virus, bovine-leukosis virus, *Mycobacterium avium* subspecies *paratuberculosis*, and *Neospora caninum* in dairy herds in the Canadian Maritimes. Prev. Vet. Med. 55, 57-68

Cohen, N.D., Lester, G., Sanchez, L.C., Merritt, A.M. and Roussel, A.J. (2004). Evaluation of risk factors associated with development of postoperative ileus in horses. J. Am. Vet. Med. Ass. 225, 1070-1078

Egenvall, A., Penell, J. C., Bonnett, B. N., Olson, P. and Pringle, J. (2005). Morbidity of Swedish horses insured for veterinary care between1997 and 2000: variations with age, sex, breed and location. Vet. Rec. 157, 436-443

Egenvall, A., Penell, J. C., Bonnett, B.N., Olson, P. and Pringle, J. (2006a). Mortality of Swedish horses with complete life insurance between 1997 and 2000- variations with sex, age, breed and diagnosis. Vet. Rec. 158, 397-406

Egenvall, A., Bonnett, B.N., Olson, P., Penell, J. and Emanuelson, U. (2006b). The association between costly veterinary care and 5-year survival of Swedish insured warmblooded riding horses. Prev. Vet. Med. 77, 122-136

Egenvall, A., Bonnett, B.N., Wattle, O. and Emanuelson, U. (2008). Veterinary-care events and costs over a 5-year follow-up period for warmblooded riding horses with and without previously recorded locomotor problems in Sweden. Prev. Vet. Med. 83, 130-143

Ekstrand, B. and Carpenter, T.E. (1998). Using a tobit regression model to analyse risk factors for foot-pad dermatitis in commercially grown broilers. Prev. Vet. Med. 37, 219-228

Kennedy, P. (2003). A guide to econometrics. The MIT press, Cambridge. 623p

Lloyd, J.W. and Kaneene, J.B. (1997). Economics of health management in the Michigan, USA equine industry. Prev. Vet. Med. 30, 1-8

Mason, B.J., Newton, J.R., Payne, R.J. and Pilsworth, R.C. (2005). Costs and complications of equine castration: a UK practice-based study comparing "standing nonsutured" and "recumbent sutured" techniques. Eq. Vet. J. 37, 468-47

Schnabel, L.V., Bramlage, L.R., Mohammed, H.O., Embertson, R.M., Ruggles, A.J. and Hopper, S.A. (2007). Racing performance after arthroscopic removal of apical sesamoid fracture fragments in Thoroughbred horses age < 2 years: 151 cases (1989--2002). Eq.Vet. J. 39, 64-68

Tobin, J. (1958). Estimation of relationships for limited dependant variables. Econometrica 26, 24-36

Traub-Dargatz, J.L., Kopral, C.A., Hillberg Seitzinger, A., Garber, L.P., Forde, K. and White, N.A. (2001). Estimate of the national incidence of and operation-level risk factors for colic among horses in the United States, spring 1998 to spring 1999. J. Am. Vet. Med. Ass. 219, 67-71

BLUETONGUE

BLUETONGUE SURVEILLANCE IN SWITZERLAND: FROM RESEARCH TO POLICY

V. RACLOZ, D. HADORN[1], H.P. SCHWERMER, C. GRIOT, A. CAGIENARD, AND K.D.C. STÄRK

SUMMARY

Due to the emergence of bluetongue disease in Europe in the last decade, the Federal Veterinary Office of Switzerland funded a series of research projects. The objective of these projects was to inform policy development. Based on the accumulated understanding of the risk factors for bluetongue (BT) in Switzerland and the changed bluetongue situation in neighbouring countries, the surveillance objectives were adjusted and consequentially, the design moved from a baseline random survey to a sentinel surveillance network and finally to nationwide risk-based sampling. This paper illustrates the use of a scenario tree model for Bluetongue virus serotype 8 (BTV-8) and how local vector population data, vector abundance along with climatic and geographical information were collected for incorporation into the model in order to make it specific for Switzerland.

INTRODUCTION

Bluetongue (BT) disease is a vector-borne viral disease of ruminant animal species which, until recently, was restricted to tropical and subtropical areas of the world. Since 1998 with its introduction into Greece, it has also been causing outbreaks in Europe with the most recent epidemic in 2006 affecting Germany, the Netherlands, France, Luxembourg, Belgium and as of September and October 2007 it has also spread into the United Kingdom, Denmark and Switzerland.

Due to the emergence of bluetongue disease in Europe in the last decade, the Federal Veterinary Office of Switzerland funded a series of research projects. The objective of these projects was to inform policy development. The first project consisted of a nationwide baseline survey conducted in 2003 and resulted in the serological screening of over 2000 cattle which confirmed the absence of BT virus infection in Switzerland (Cagienard et al., 2006). An entomological study conducted at the same time and focusing on southern, western and south-eastern parts of the country resulted in the identification of potential BT vectors in many areas. This study prompted the initiation of another project with the aim of establishing an early warning system for BT in Switzerland and was conducted from 2004-2007 (Racloz et al., 2006). Serological and entomological surveillance was carried out during this period at 14 sentinel sites, along with the collection of climatic and geographical data. Through this data collection, thematic maps were created in order to determine areas most suitable for vector activity. Using spatial and temporal methods, it was possible to establish the areas and periods of the year which are most important for the risk of incursion and spread of BT disease.

[1]Daniela Hadorn, Swiss Federal Veterinary Office, Schwarzenburgstrasse 155, 3003 Bern, Switzerland. Email: daniela.hadorn@bvet.admin.ch

Based on the accumulated understanding about the risk factors for BT in Switzerland and the changed BT situation in neighbouring countries, the surveillance objectives were adjusted and consequentially, the design moved from a sentinel surveillance network that was relatively limited in farm numbers to a nationwide risk-based sampling strategy. The objective was to design an optimal surveillance system which maintained a risk-based element and covered the whole country, yet remained within a pre-defined financial budget, considering that the country was still free from BT at that time. To reach this goal, a stochastic scenario tree modelling method was used (Hadorn & Stärk, submitted; Martin et al., 2007) in order to determine what type of surveillance system components (SSC) were best suited in Switzerland for a national surveillance system of BT in terms of sensitivity performance and cost.

This paper illustrates the use of a scenario tree model for bluetongue virus serotype 8 (BTV-8) and how local vector population data, vector abundance along with climatic and geographical information were collected for incorporation into the model in order to adapt it to Switzerland's specific conditions.

During the time of writing, Switzerland reported five outbreaks of BT (www.bvet.admin.ch). These were located in areas considered at high risk of BT occurrence based on vector biology. On the first three farms, BT was identified through a clinical suspect case followed by serological testing. Out of a total of 240 susceptible cattle from the farms located in Basel-Stadt, Solothurn and Basel-Land, it was detected in nine animals, with prevalences of 10.3%, 2.4% and 1%, respectively. Recently, through bulk milk testing, a fourth outbreak affecting a single cow in a herd of 80 cattle was reported in the canton of Basel-Land, and the most recent outbreak was reported in goats in the canton of Solothurn.

MATERIALS AND METHODS

The scenario tree methodology allows assessment and comparison of the outcomes of a variety of interventions, in this case SSC, for example by varying the input values (i.e. new data on a monthly basis). As described by Martin et al. (2007), five main steps need to be followed in order to construct a scenario tree. Firstly, to determine the order of events affecting the objective of the scenario tree. Secondly to include the livestock structure of the country or area involved in the model, followed by identifying the risk factors involved in the disease/condition. The fourth step is to incorporate the testing and sampling methods used, and finally to assess the feasibility of the program (www.ausvet.com.au/freedom).

In this study, the objective of the scenario tree was to determine what SSC could be used as an effective and economical program to detect BT in Switzerland. The livestock involved in this model were cattle and sheep populations. The most complex steps were to determine what risk factors were involved in the case of BT infection, along with analyzing various detection methods for each SSC as described below.

Risk factors for BT integrated in the stochastic simulation model

Bluetongue disease is a non-contagious, infectious, insect-transmitted viral disease that affects domestic and wild ruminants (Purse et al., 2005), with its occurrence exclusively related to the presence of competent vectors. Therefore, the main risk factors for the presence of BT infection are areas of the country which have suitable climatic, geographical, host associated and entomological features for the sustainment and spread of the virus.

Spatial risk factors

Climatic and geographic areas suitable for the establishment of the vector were determined through the creation of monthly thematic maps for altitude, precipitation and average temperatures using ArcGis (Version 8.3 , Environmental Systems Research Institute, Inc.), and data from 50 meteorological stations provided by the Swiss Meteorological Office for the year 2006 (Racloz et al., 2007). Smoothed maps were generated using ordinary kriging apart from the altitude map which was derived from an existing elevation model, then combined using the raster map calculator function for producing individual monthly suitability maps for 2006. Suitability categories were generated on the basis of a review of the literature available on *Culicoides* biology (Mellor et al., 2000, De Liberato et al., 2005, Purse et al., 2005, Carpenter et al., 2006, Osmani et al., 2006, Purse, 2006), and subsequently divided into four categories, ranging from high to low.

Transmission risk based on R_0 calculations

The geographic risk areas were used to determine the relative risk of disease transmission by calculating the basic reproduction number (R_0) for each suitability risk zone. For vector-borne diseases, R_0 is considered as the number of secondary cases a single infected vector will produce in a susceptible population of hosts (Gubbins et al., 2007). This was used to attribute a transmission level to each geographical risk category in order to determine the efficacy of risk based sampling. Hence, each monthly suitability map would display different areas of high to low risk zones, and an R_0 value was calculated using entomological data collected through the sentinel herd surveillance program in Switzerland (2004-2007) (Racloz et al., 2006b). The R_0 equation used was based on previous malaria (Smith et al., 2007b) and West Nile models (Wonham et al., 2004), as well as current knowledge of BT epidemiology (Gubbins et al., 2007). This allowed visualizing the consequence of an outbreak depending on its geographical starting point (Racloz et al., submitted).

Surveillance system components (SSC)

The second step was to determine what possible surveillance methods were available for the BT and which populations had to be surveyed. As discussed in Hadorn & Stärk (submitted), a surveillance system may be composed of both active and passive surveillance parts. The ability to combine various independent SSCs (Martin et al., 2007) and to estimate an overall detection probability for the surveillance system all within an economically viable process is an important feature of scenario tree modelling, especially for emerging disease surveillance. Since the aim of this project was to establish a federal surveillance program, it was important to include criteria specified by the OIE concerning BT surveillance. Hence the model was designed to assess the surveillance system for BT assuming a design prevalence of 0.2%.

The following potential SSCs were identified for BT surveillance in both cattle and sheep populations: serological random sampling of cattle and sheep, randomly selected bulk milk testing of dairy herds, abortion reporting and testing in cattle and sheep, isolation of virus from vectors, serological slaughterhouse sampling in cattle and sheep, clinical surveillance in cattle and sheep, and finally risk-based sentinel herd sampling (serological and bulk milk testing). After creating a basic scenario tree with these SSCs, the respective component sensitivities and their economic implications were evaluated (data not shown). For economic and practical reasons, the following three SSCs were retained for further analysis: passive surveillance

strategy in terms of clinical surveillance of sheep and cattle, as well as the active SSC of bulk milk testing in sentinel herds.

Sentinel Herd Bulk Milk testing in dairy cattle

Sentinel herd bulk milk testing means that a certain number of herds with an increased risk of getting BT-infection are tested monthly using the milk test ELISA (ID Screen® Blue Tongue Milk from ID Vet, France). In Switzerland, a regular nation-wide milk sampling procedure already exists as part of the milk quality testing program and the bulk milk samples for the BT surveillance program could be integrated into this program. Therefore, costs for bulk milk sampling proved to be much lower than a similar surveillance method based on serological blood sampling for individual animal testing. According to the analysis of the basic model with respect to costs and system sensitivity benefit, it was decided that a maximum of 200 herds could be integrated in this program. Herds were to be located in areas considered to be at high BT and vector risk.

The input parameters used to determine the overall detection probability (Se) for this SSC were 1) the risk factors, including the distribution of the geographic risk areas as well as the relative risk for vector activity levels calculated using the basic reproduction number (R_0), and 2) the detection probability of positive herds through bulk milk testing using the commercially available ELISA. For this SSC, only the herds distributed in the high and hi-medium risk categories were eligible.

Clinical surveillance in cattle and sheep

Clinical surveillance (CLIN) consists of the detection and reporting of suspect infected farms and animals through animal owners, caretakers or veterinarians. The key elements in this process are the probability that infected animals show clinical symptoms as well as the disease awareness (DA) in farmers and veterinarians.

Due to the nature of the northern European outbreak and the exclusive involvement of BT serotype 8 (BTV-8), it was possible to collect data on clinical symptoms from affected countries and use it as an input parameter for this SSC. The disease awareness levels of farmers and veterinarians, i.e. the probabilities of the farmer contacting the veterinarian, and that of the veterinarian conducting the appropriate BT diagnostic test, which involves serological blood sampling using an ELISA (VMRD Inc., Pullman, WA, USA), were set arbitrarily using similar disease awareness categories as described in Hadorn & Stärk (submitted). In the case of CLIN for cattle, the DA was increased by 5% according to previous experience on animal health disease awareness. The assumption was that cattle husbandry was more focused on individual animals than sheep husbandry and therefore the probability to consult a veterinarian in case of clinical symptoms in single cattle is slightly higher than for a single sheep. Because BT symptoms were reported to be more common in sheep than in cattle (Mehlhorn et al., 2007), the DA levels of veterinarians to take samples for BT in suspect sheep were set 5% higher than in suspect cattle.

A low DA was attributed to CLIN in cattle, from January to June, and a low-medium DA from July to December because the increase in BT cases in Northern Europe was expected to impact DA, as well as it being the vector activity period. For CLIN in sheep, input parameters of DA were modified depending on the month during which the SCC were active. Low DA was used for the months of January to July, a low-medium DA level for August and September,

ending with the medium-high DA levels for October to December. The differences in the DA values were justified on the basis of media output of the Federal Veterinary Office, as well as the pattern of the vector season and finally due to the BTV-8 situation in northern Europe. Values for the different DA levels are provided in Table 1.

In this CLIN-SSC, all geographic risk zones were taken into consideration, and the final detection probability level for clinical surveillance in cattle and sheep was calculated considering the total number of cattle herds (37,860) and sheep (22,201) in Switzerland.

Input parameters

Input parameters which needed to be calculated for every month included the following: proportion of herds in the different geographic risk categories, vector activity rates to determine the relative risk of each geographic zone, disease awareness (DA) estimates, clinical symptom data and diagnostic test performance. The final step was to incorporate host distribution data and to determine the percentage of herds in each risk and suitability category on a monthly basis, also taking into consideration the altered distribution of cattle herds during summer due to the alpine pasture tradition. As mentioned before, all risk factor values were calculated on a monthly basis due to the fast moving nature of BT, as well as the effect of different climatic patterns.

Values for these parameters were calculated either as fixed values as in the case of herd distribution or as Pert distribution values. The Pert distributions accounted for the uncertainty in the data, and permitted a range for minimum, most likely and maximum values to be calculated when running the Monte Carlo simulation.

Once all the input parameters and the risk factors had been determined for each month, separate monthly Monte-Carlo simulations were run for each SSC by using the @Risk software program (Palisade Corporation) for 5,000 iterations. The overall combined detection probability output was then calculated once the SSC values for all months were determined.

The three SSCs, namely the sentinel herd bulk milk testing and the clinical surveillance in cattle and sheep, were combined to obtain a final overall value for the performance of the entire surveillance system as described by Hadorn & Stärk (submitted) and Martin et al. (2007). Because BT is a fast moving disease, our time step in analysis was one month.

RESULTS

The performance of the selected SSCs can be compared in Table 1. After exploring several combination options (data not shown), a policy was developed which uses the Bulk Milk Testing SSC, the clinical surveillance in cattle along with the clinical surveillance in sheep, which produced the highest combined Se levels, along with an acceptable economic output in terms of the Swiss BT situation at that time.

Changes in the risk factors for BT were seen on a monthly basis in both the geographic distribution risk factor and the relative risk based on R_0 calculations as discussed in Racloz et al. (submitted). The peak of the highest geographic risk occurred in the month of September, followed by August, July and May. In terms of the relative risk values, June represented the highest risk, followed by July and September. These two input parameters had the largest effect on the detection probability of the sentinel herd bulk milk testing SSC component, which can be seen in the fluctuations of the Se and the R_0 values.

The detection probability values for the chosen SSCs ranged from 0% to 80% over the year, but reached the highest levels in the second half of 2007. Overall, the CLIN SSC for cattle had the highest Se levels and a range of 35%-80%, followed by the sheep CLIN SSC which had values from 17%- 78% whilst the sentinel bulk milk testing SSC had Se levels ranging from 0%-41%.

The overall sensitivities along with the individual output sensitivities of the three SSCs are shown in Fig. 1. Disregarding the sentinel surveillance system which was in place prior to 2007, the combined Se of the passive clinical surveillance both in sheep and cattle was considered to be 46% until the month of June. At this time point, the bulk milk testing program was put into place and raised the combined Se to above 90%. The fluctuation seen in the bulk milk testing SSC is due to the climatic effect on the R_0 values which in turn affect the probability of detecting BT as a result of varying vector activity rates.

Table 1. Annual probability of case detection using individual or combined surveillance system components for bluetongue surveillance in Switzerland

Surveillance system component	System detection probability
Abortion testing	0.3%
Sentinel Herd Surveillance (SHS) Bulk milk testing	38.7%
Passive clinical surveillance in sheep	62.4%
Passive clinical surveillance in cattle	85.9%
Combined SHS and Clinical surveillance for cattle and sheep	97.4%

DISCUSSION

The results of a series of research projects, including a scenario tree model enabled decision makers to establish an evidence-based national surveillance plan for BT disease in Switzerland. The results of the projects provided information on the initial infection status of livestock and then informed the development of risk-based surveillance, first using sentinel herds and then applying nation-wide testing. The model highlighted the importance of disease awareness and its effect on detection probabilities. Subsequently, various workshops and meetings were organized involving stakeholders and animals holders, especially sheep farmers. Direct results could be seen with the amount of clinical suspect cases that were being reported in comparison to previous years. We also showed that the sentinel bulk milk testing SSC, although not as sensitive as a blood serological screening, allowed the overall Se of the program to reach adequate levels of effectiveness within the limits set by the financial budget. From July 2007 onwards, regular bulk milk testing was carried out which also resulted in three false positive results. In these cases, serological blood testing of the whole herd was carried out. A recent bulk milk sample from the canton of Basel-Land tested positive in mid-November 2007, and subsequent serological blood sampling resulted in one BT positive cattle, making it the fourth case reported in Switzerland. This demonstrated that the bulk milk test was able to pick up positive signals when only one out of 80 animals contributing to the bulk tank was infected.

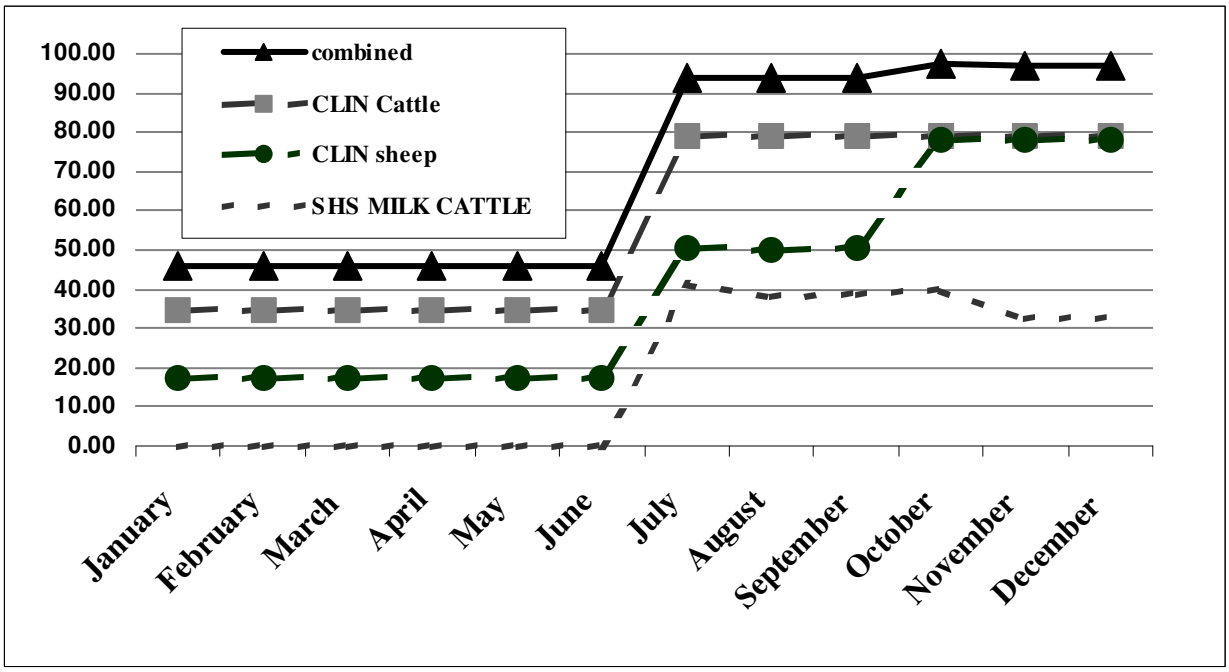


Fig. 1 Probability of case detection using individual or combined surveillance system components for bluetongue surveillance in Switzerland by month. CLIN = clinical SSC, SHS= Sentinel Herd Bulk Milk testing SSC.

The main increase in detection levels for the clinical surveillance in cattle SSC occurred in late July due to the estimated increase in DA related to the release and distribution of a BT documentary movie on DVD to stakeholders, and the increased number of reports in Swiss newspapers.

The SSC for sheep had two main increases in detection probability, whereby the first increase being due to the activity mentioned above, along with a second increase in DA resulting from the education program aimed at a selected number of sheep farmers and union members.

The results highlight the importance of input parameter quality and their effect on the overall performance of the surveillance system. The input parameters related to DA levels were highly influential at improving the overall performance of the surveillance program. In previous models (data not shown), various simulations were carried out with either a combination of DA levels, or running the whole model at a single DA level. Modest DA levels were used in the final simulation in order to generate conservative model results. In the future it is planned to update the model with more precise data based on indicators such as the level of media interest, numbers of suspect cases being reported, number of actual bulk milk testing samples, real data on veterinarian and farmer contacts and correct diagnostic procedures, as well as the amount of enquiries reaching the Federal Veterinary Office from the public concerning BT disease.

Another important finding was the impact of assumptions regarding vector activity on the Se of the sentinel bulk milk testing component. The fluctuations seen emphasized the complexity of developing a surveillance system for vector-borne diseases. Environmental and transhumance factors, along with climatic fluctuations will affect the activity and survival of the vector, and in turn alter the overall surveillance sensitivity by lowering or increasing the detection levels. As input parameters remain uncertain, numeric results of the simulation need to

be interpreted with care. However, the relative comparison of SSC performance is considered to be sufficiently robust to be used as a basis for decision making.

Each country uses different methods for BT surveillance depending on the current disease status, geographical barriers, allocated finances and resources available. Most countries involved in the current northern European outbreak have SSCs similar to Switzerland, although serological blood sampling is used more widely. However, our model indicated that this strategy did not significantly increase the overall Se in Switzerland, while being substantially more expensive.

Switzerland detected its first BT case in late October 2007 through clinical surveillance. During the following six months, reporting of suspect cases increased from a total of five suspect cases in 2006, to 33 cases being reported before the first outbreak in 2007 (B. Thür, personal communication), which is likely to indicate increased disease awareness.

In conclusion, through the analysis of various BT disease surveillance options, three SSCs were included in a national program for the early detection of BT in Switzerland. By identifying the most influential input parameters, actions were taken to strengthen these parts of the surveillance system. This included the production of a BT documentary movie of which more than 3000 DVD copies have already been distributed. Additionally, workshops on the disease have been delivered in different regions of the country in order to raise disease awareness, especially in the sheep industry. This system illustrates how surveillance results can inform risk assessment and risk assessment results in turn can be used to target surveillance efforts resulting in a risk-based surveillance system. It was also shown how results can help identifying and targeting weaker areas of disease awareness and information distribution.

ACKNOWLEDGEMENTS

The authors would like to thank Dr. Eric Breidenbach, Dr. Hansueli Ochs and Dr. Lukas Perler for their valuable input.

REFERENCES

Cagienard, A., Thur, B., Griot, C., Hamblin, C., and Stark, K.D. (2006). No evidence of bluetongue virus in Switzerland. Vet. Microbiol. 116, 13-20

Carpenter, S., Lunt, H.L., Arav, D., Venter, G.J. and Mellor, P.S. (2006). Oral susceptibility to bluetongue virus of Culicoides (Diptera: Ceratopogonidae) from the United Kingdom. J. Med. Entomol. 43, 73-78

De Liberato, C., Scavia, G., Lorenzetti, R., Scaramozzino, P., Amaddeo, D., Cardeti, G., Scicluna, M., Ferrari, G. and Autorino, G.L. (2005). Identification of *Culicoides obsoletus* (Diptera: Ceratopogonidae) as a vector of bluetongue virus in central Italy. Vet. Rec. 156, 301-304

Gubbins, S., Carpenter, S., Baylis, M., Wood, J.L. and Mellor, P.S. (2008). Assessing the risk of bluetongue to UK livestock: uncertainty and sensitivity analyses of a temperature-dependent model for the basic reproduction number. J. R. Soc. Interface 5(20), 363-371

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

Mehlhorn H. Walldorf, V., Klimpel, S., Jahn, B., Jaeger, F., Eschweiler, J., Hoffmann, B. and Beer, M. (2007), First occurrence of *Culicoides obsoletus*-transmitted bluetongue virus epidemic in central Europe. Parasitol. Res. 101, 219-228

Mellor, P.S., Boorman, J. and Baylis, M. (2000). Culicoides biting midges: their role as arbovirus vectors. Annu. Rev. Entomol. 45, 307-340

Osmani, A., Murati, B., Kabashi, Q., Goga, I., Berisha, B., Wilsmore, A.J. and Hamblin, C. (2006). Evidence for the presence of bluetongue virus in Kosovo between 2001 and 2004. Vet. Rec. 158, 393-396

Purse, B.V., Baylis M., McCormick B.J.J. and Rogers D. (2006). Hindsight and foresight on the spread of bluetongue virus in Europe. Report T.8.3 as part of Foresight project 'Infectious Diseases: preparing for the future'. Office of Science and Innovation, UK. 46pp

Purse, B.V., Mellor, P.S., Rogers, D.J., Samuel, A.R., Mertens, P.P. and Baylis, M. (2005). Climate change and the recent emergence of bluetongue in Europe. Nat. Rev. Microbiol. 3, 171-181

Racloz, V., Presi, P., Vounatsou, P., Schwermer, H., Casati, S., Vanzetti, T., Griot, C., and Stärk, K.D.C. (2007). Use of mapping and statistical modeling for the prediction of Bluetongue occurrence in Switzerland based on vector biology. Veterinaria Italiana 43(3), 513-519

Racloz, V., Straver, R., Kuhn, M., Thur, B., Vanzetti, T., Stark, K.D., Griot, C., and Cagienard, A., (2006). Establishment of an early warning system against bluetongue virus in Switzerland. Schweiz. Arch. Tierheilkd. 148, 593-598

Smith, D.L., McKenzie, F.E., Snow, R.W. and Hay, S.I. (2007). Revisiting the basic reproductive number for malaria and its implications for malaria control. PLoS biology 5(3), e42

Wonham, M.J., de-Camino-Beck, T. and Lewis, M.A. (2004). Anepidemiological model for West Nile virus: invasion analysis and control applications. Proc. Biol. Sci. 271, 501-507SA

ESTABLISHING THE SPREAD OF BLUETONGUE VIRUS DURING THE 2006 EPIDEMIC IN BELGIUM

E. MEROC*, C. FAES, C. HERR, B. VERHEYDEN, T. VANBINST, F. VANDENBUSSCHE, J. HOOYBERGHS, M. AERTS, K. DE CLERCQ, AND. MINTIENS

SUMMARY

In August 2006, Bluetongue (BT) was notified for the first time in Northern Europe. In Belgium, during the epidemic, case-reporting relied almost exclusively on the identification of herds with confirmed clinical infected ruminants. A cross-sectional serological survey targeting all Belgian cattle was then undertaken. The study's objectives were to: 1. provide unbiased estimates of BT-seroprevalence; 2. compare final dispersion of the virus based on the seroprevalence estimates to that of case herds. True within-herd seroprevalence was estimated based on a logistic-normal regression model with prior specification of the diagnostic test's sensitivity and specificity. Herd seroprevalence was estimated using a logistic regression model. To study the linear correlation between serological survey and case-herd data, the linear predicted values for the herd seroprevalences were compared and the Pearson correlation coefficient was estimated. Overall herd and true within-herd seroprevalences were estimated at 83.3% (79.2-87.0) and 23.8% (20.1-28.1), respectively. The analysis has shown there was a strong correlation between the two datasets ($r=0.73$, $p<0.001$). The case detection system underestimated the real impact of the epidemic, but provided an accurate indication of the spatial distribution of the infection.

INTRODUCTION

Bluetongue (BT), a vector-borne viral disease, is transmitted in ruminant populations almost exclusively by several species of biting midges of the genus *Culicoides* (Diptera: Ceratopogonidae) (Gibbs & Greiner, 1994). BT virus (BTV) is a species of the genus *Orbivirus* within the *Reoviridae* family. To date, 24 distinct BTV-serotypes have been identified. BT can cause spectacular outbreaks and has an adverse impact on worldwide trade due to restrictions in relation to the source of animals (FAO, 2006). It thus appears on the list of diseases notifiable to the World Organisation for Animal Health (OIE). Influenced by several factors such as geographical location, the incidence of clinical disease is highly variable. BT disease is uncommon in many areas where BTV is endemic (MacLachlan, 2004). The virus is traditionally known to be distributed around the world in countries located in the tropics and subtropics, although it may extend further north such as in parts of western North America and Xinjiang, China (Dulac et al., 1989; Gibbs et al., 1994; Qin et al., 1996). The virus has recently been documented as far as 45°N in southern Europe (Caporale et al., 2004).

* Estelle Méroc, Veterinary and Agrochemical Research Centre, Co-ordination Centre for Veterinary Diagnostics, Groeselenberg 99, 1180 Brussels, Belgium. Email: estelle.meroc@var.fgov.be

In August 2006, very unexpectedly, BT was for the first time reported from the Netherlands, Belgium and Germany (OIE Animal Health Department, 2006). Later during the epidemic, related cases were also reported from France and Luxembourg. The virus incriminated was identified as BTV-serotype 8 (CRL, 2006; Toussaint et al., 2007a), which prior to this epidemic had only occurred in Africa, Central America, Malaysia, and India/Pakistan (Herniman et al., 1980; Hassan, 1992; Mo et al., 1994; Daniels et al., 2004; Gerdes, 2004). Based on the data from the early stages of the epidemic, the rate of local spread was estimated to be around 15 km/week, partially reflecting the rapid extension of BTV in northern Europe (Gerbier et al., 2007).

In Belgium, the first 11 BT outbreaks reported were confirmed in the part of the country near the eastern border on the 19th of August 2006, in both sheep and cattle herds (Toussaint et al., 2007b). Despite the implementation of an animal movement ban, the disease was rapidly and widely disseminated throughout the Belgian territory. By December 2006, a total of 695 herds or flocks were declared "case herds" of which 297 were cattle herds. During the epidemic, case-reporting by the Belgian Veterinary Authorities relied almost exclusively on the identification of herds with confirmed clinical infected ruminants. Laboratory diagnoses were mostly used for confirmation of BTV infections in ruminants reported with BT-like clinical signs. Therefore, under-reporting was suspected.

During the winter of 2006/2007, it was assumed that climatic conditions were unfavourable for further propagation of BTV. The last cases of the epidemic in Belgium were reported by the Veterinary Authorities on 15 December 2006. A serological and virological cross-sectional survey (BT winter screening) targeting all Belgian cattle was undertaken in January-February 2007 in order to establish the true final dispersion of the virus across the country. The first objective of the study was to provide unbiased estimates of BT-seroprevalence for different regions of Belgium. A second objective was to compare the final dispersion of the virus based on the seroprevalence estimates to the dispersion of the confirmed clinical cases which were notified in Belgium, in order to estimate the accuracy of the case-detection based on clinical surveillance. This paper presents the descriptive epidemiology of the BT winter screening 2007.

MATERIALS AND METHODS

Sampling design for the BT winter screening

The study population of the winter screening consisted of dairy cattle more than two years old which were housed in dairy farms with on-farm delivery of dairy products. Only dairy cattle were considered for sampling since serologically negative animals that were to be identified by the BT winter screening would participate subsequently in a longitudinal BT sentinel animal monitoring programme (dairy animals are sampled more easily). The sampling frame was based on the 1245 dairy herds with on-farm delivery of dairy products previously identified for the official Belgian Leucosis-Brucellosis winter screening. As part of this programme, all animals of more than two years were sampled. Since no prior information on the herd prevalence was available, the number of herds to be sampled was based on an expected prevalence of 50% (maximal variance), a desired absolute precision of 5% and 95% confidence level. Since the diagnostic test was assumed to have perfect sensitivity and specificity, a sample of 384 herds was to be selected (Cannon .and Roe, 1982). A one-stage sampling design was used with stratification of the herds by province and proportional allocation according to province surface area.

Diagnostic methods

Samples were collected by the official farm veterinarians and serum samples were extracted at the regional laboratories of 'Dierengezondheidszorg Vlaanderen' and the 'Association Régionale de Santé et d'Identification Animales'. The serum samples were assayed using a commercially available competitive ELISA (c-ELISA) kit (ID Screen® Blue Tongue Competition for detection of anti-VP7 antibodies; ID.VET, Montpellier, France) which was carried out according to the OIE Manual of Standards (OIE, 2004) and to the procedure described by the manufacturer. Results were expressed as percentage negativity (PN) compared to the negative kit control and cut-off settings considered were those provided by the manufacturer. Samples which presented a PN less or equal to 35%, between 35 and 45%, and greater than 45% were considered as positive, doubtful and negative, respectively. Doubtful results were classified as positive in the data analysis. Using RT-qPCR as reference test during the epidemic, the diagnostic sensitivity and specificity of the c-ELISA was estimated at 87.4% (95%CI: 83.5-90.4) and 99.0% (95%CI: 97.2-99.6), respectively (Vandenbussche et al., 2007).

Case herds

Case herds were mostly herds (cattle or ovine) for which the veterinary practitioner, who had been consulted by the animal owner, identified suspicious clinical cases and where at least one of those animals was subsequently confirmed positive using a laboratory test (c-ELISA and/or real-time PCR) and then notified to the veterinary authorities (EFSA, 2007). A maximum of three animals were sampled per herd. In addition, herds without clinical signs but with seropositive animals which were then confirmed positive with real-time PCR (Toussaint et al., 2007a) were also included. For example, animals could be detected when tested serologically for certification prior to trade between zones with different BTV-8 status within the country or prior to export. EDTA blood and serum samples were tested at the Belgian National Reference Laboratory (VAR).

Statistical methods

For the BT winter screening data, let Z_i be the number of positive tested animals out of N_i tested animals from herd *i*. Further, let $Y_i = 1$ when at least one animal in a herd *i* tested positive, and *0* otherwise.

Estimation of the within-herd seroprevalence was based on a logistic-normal regression model. It is assumed that the number of positive animals follows a binomial distribution:

$$Z_i \sim Bin(N_i, p_i^a) \tag{1}$$

with p_i^a the apparent seroprevalence. However, since interest is in the true seroprevalence p_i^t, reflecting the true disease status of the animals, the sensitivity and specificity of the ELISA test should be accounted for. From the apparent seroprevalence p_i^a, the true seroprevalence p_i^t can be derived from the following equation (Rogan and Gladen,1978):

$$p_i^t = \frac{p_i^a + Sp - 1}{Se + Sp - 1} \tag{2}$$

where *Se* is the test sensitivity and *Sp* is the test specificity. To account for possible correlation among the animals from the same herd, the prevalence of disease in herd *i* is modelled as:

$$\text{logit}(p_i^t) = \beta + u_i \tag{3}$$

with normally distributed random intercepts for each herd $u_i \sim Normal(0, \sigma^2)$. This is a special form of a generalized linear mixed model as described by Molenberghs and Verbeke (2005). Since the sensitivity and specificity are no fixed or known values, a prior distribution for the sensitivity and specificity was assumed. Thus, model specification is further extended by assuming a beta-distribution for the *Se* and *Sp* parameters:

$$\text{Se} \sim \text{beta}(a_1, b_1) \tag{4}$$

$$\text{Sp} \sim \text{beta}(a_2, b_2) \tag{5}$$

where a_1, b_1, a_2, b_2 are chosen based on published data. The model is given by Eq. (1) to (5). Because of its hierarchical structure, it is fitted in a Bayesian framework, using the WinBUGS software (http://www.mrc-bsu.cam.ac.uk/bugs). Non-informative priors were used for all model parameters. Posterior seroprevalence distributions and 95% credibility intervals were generated. A map showing the distribution of within-herd seroprevalence estimates was produced using ArcView GIS 3.2. (ESRI).The estimated true prevalence values of the farms which were sampled were interpolated using inverse distance weighting interpolation based on the 6 nearest neighbouring farms.

Herd seroprevalence (probability that a herd was infected) was estimated using a logistic regression model:

$$Y_i \sim Bernouilli(p^h) \tag{6}$$

$$\text{Logit}(p^h) = \beta \tag{7}$$

where p^h is the apparent herd seroprevalence. This model can be easily extended by allowing different β's for the different provinces, in order to estimate the province-specific herd seroprevalences. For the purpose of this study, a herd was considered as positive if at least one of the sampled animals had a positive c-ELISA result, otherwise it was considered negative.

In both models described above, the design-effect was taken into account by weighting each observation by the inverse of the sampling probability. Livestock density data was extracted from the Belgian animal identification and registration system (SANITEL) to provide estimates of the population at risk.

In order to estimate the accuracy of the case-detection based on clinical surveillance, the linear correlation between the BT winter-screening data and the case data was estimated. For both data sets the herd seroprevalence per municipality was estimated based on logistic

regression models, as in Eq. (6). In order to account for spatial differences, a flexible smoothing method was used to estimate the spatial trend. It is assumed that:

$$\text{Logit}(p^h) = f(x, y) \tag{8}$$

with $f(x, y)$ being an unspecified smooth function of the *x*- and *y*-coordinates. The method focused on, was the use of penalized splines with radial basis function, fitted as a generalized linear mixed model (Eilers & Marx, 1996). This method was implemented in the SAS procedure GLIMMIX. Finally, the linear predicted values $f(x, y)$ for herd prevalences, generated based on the logistic regression models described above for the two datasets, were compared and the Pearson correlation coefficient was estimated. For the outbreak data to be comparable with the winter screening data, solely cattle results were used for this part of the analysis.

RESULTS

Winter screening findings

A total of 25,846 cattle from 344 herds were sampled between the first and the 31st of January 2007. An average of 75 animals, ranging from 1 to 370, was sampled per herd. Among those samples, 5008 gave positive results. The overall herd seroprevalence was estimated at 83.3% (95%CI 79.2-87.0). Province-specific herd seroprevalences and their credibility intervals are shown in Fig. 1. The true overall within-herd seroprevalence was 23.8% (95%CI 20.1-28.1). The spatial distribution of the within-herd seroprevalence is presented in Fig. 2. The highest provincial within-herd seroprevalence estimates were found in Limburg and Liege. Around the city of Ghent in East Flanders, within-herd seroprevalence is also high. The estimate at provincial level is the lowest in Hainaut.

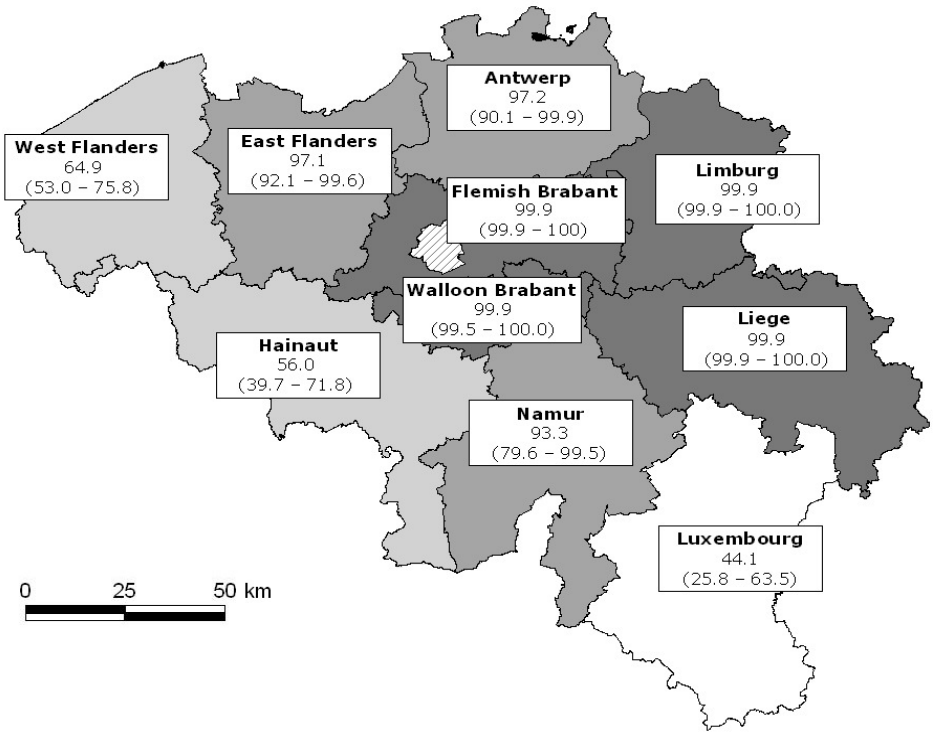


Fig. 1 Province-specific BT herd seroprevalence (in %) in Belgian dairy cattle based on the data from the winter screening, January 2007

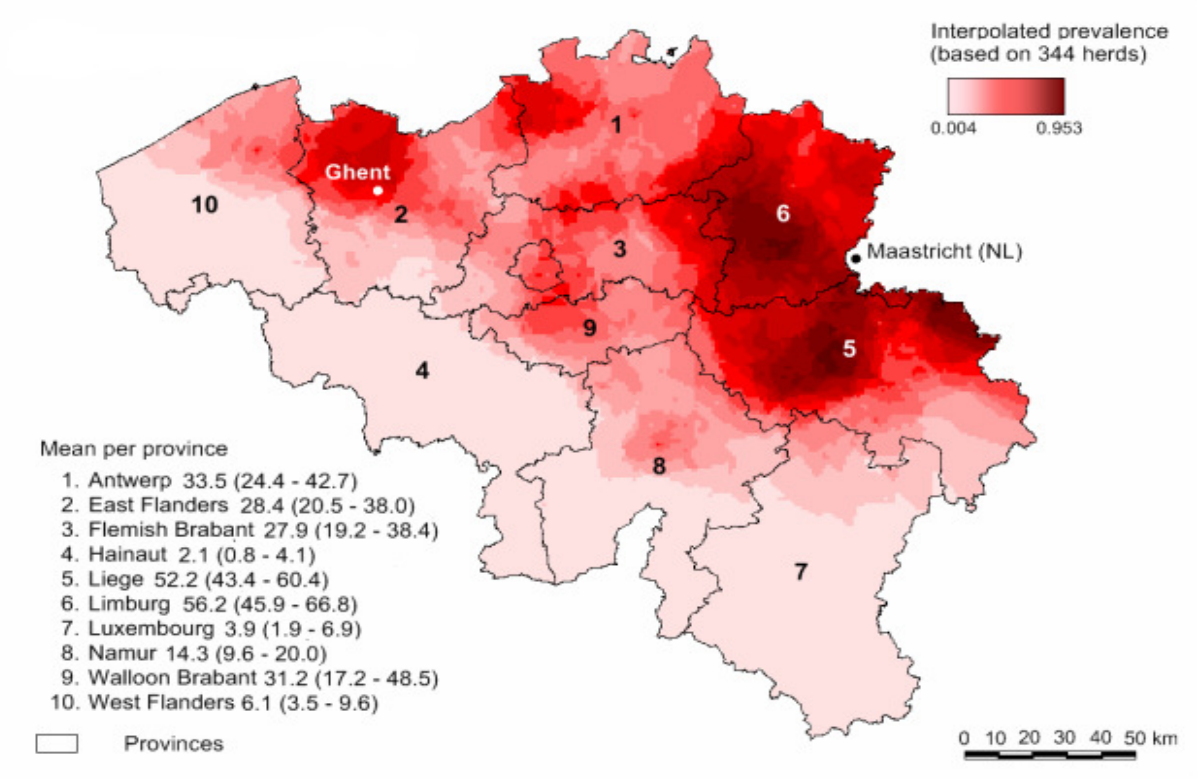


Fig.2 Distribution of within-herd BT seroprevalence (in %) in Belgian dairy cattle based on the data from the winter screening, January 2007

Case herds

Between the 18th of August and the 31st of December 2006, a total of 1445 cattle and 893 sheep samples were analysed. The overall herd-prevalence was estimated at 0.7% (95%CI 0.7-0.8) for cattle herds and 1.3 % (95%CI 1.2-1.4) for sheep herds. Herd-prevalence at provincial level is shown separately for cattle and sheep case herds respectively in Fig. 3 and Fig. 4.

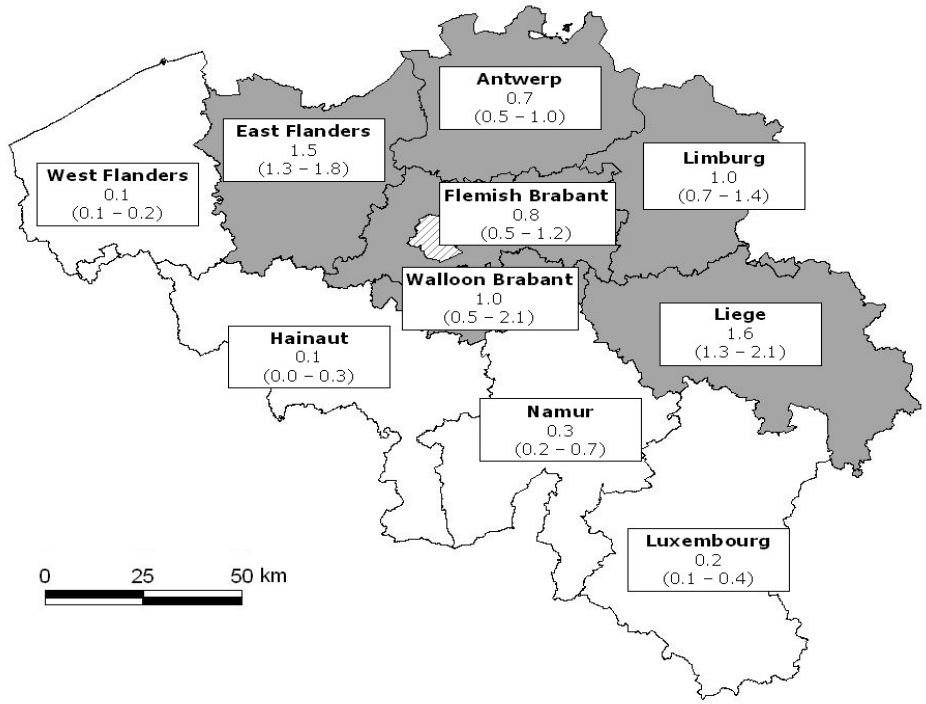


Fig.3 Province-specific BT herd-prevalence (%) in Belgian cattle herds based on the case herd data, 18 Aug- 31 Dec, 2006

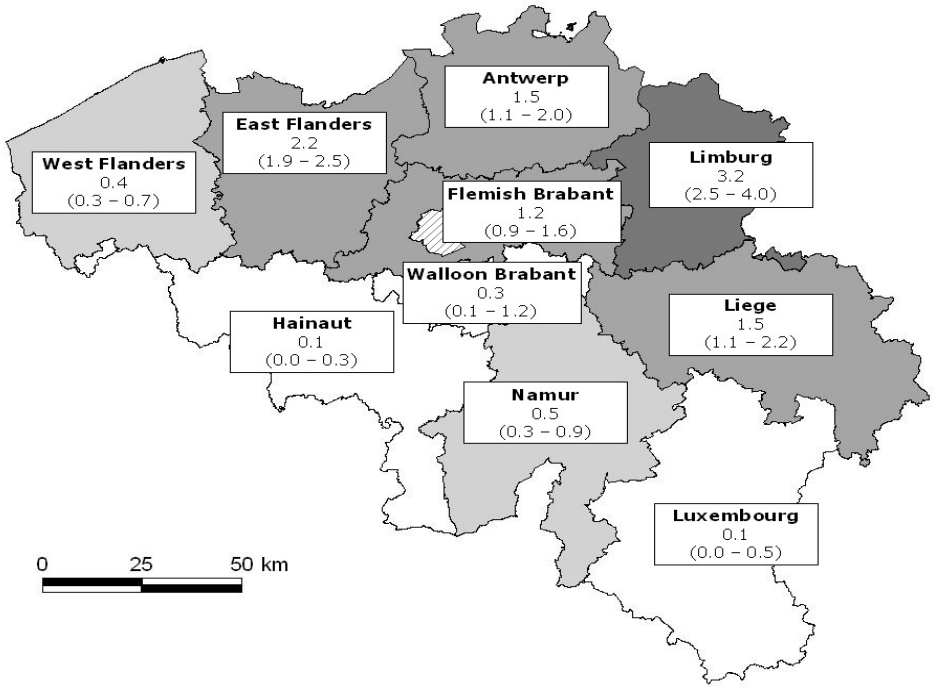


Fig.4 Province-specific BT herd-prevalence (in %) for Belgian sheep flocks base don the case herd data, 18 Aug-31 Dec, 2006.

Comparison of winter screening and cattle case herd results

Based on the spatial regression estimates for the winter screening and the cattle case herd data, the Pearson correlation coefficient was 0.73 (p-value<0.0001).

DISCUSSION

Starting from the original focus in the area where Belgium, The Netherlands and Germany share borders, the epidemic gradually disseminated throughout the northern European countries. The epidemic predominantly spread horizontally along an east-west axis. In Belgium, until October 2006, case herds were mainly limited to an area situated in the Eastern part of the country. Early September 2006, the area of main concern appeared to be the infectious status of the "still free" provinces; therefore, a serological screening was conducted and demonstrated freedom of BTV infection for all provinces in which no case herd had been notified at that time (Vandenbussche et al., 2007). The first case in East Flanders was notified on September 18 and the infection then continued its further spread towards the west. At the end of the epidemic period, BTV-seropositivity in dairy cattle herds was shown to be widely but unevenly distributed throughout Belgium. Based on case herd data, Gerbier et al. (2007) identified two spatial clusters of cases in Belgium which center around the cities of Maastricht (the Netherlands) and Ghent. The authors stated that a gap between the two clusters remained until the end of the epidemic. The results of the winter screening on the other hand (Fig. 1) showed a homogeneous herd seroprevalence between Gent and Maastricht (prevalence was found to be above 97% in Liege, Limburg, Antwerp, Flemish Brabant and East Flanders). The interpolated map showing the distribution of within-herd seroprevalence revealed areas around Maastricht and Ghent where the within-herd seroprevalence was high. In the present study, the highest within-herd seroprevalences were found on farms situated in Liege, Limburg, and Brabants, most certainly due to the fact that those regions were affected at the beginning of the epidemic. The second focus around the city of Ghent could be explained, for instance, by the high cattle farm density in this area, which could be a risk factor for within-herd propagation of BT. However, further study of specific risk factors such as local temperature, farm management system, and abundance of vector, is needed to better understand the spatial variation in the occurrence of BT and to allow more efficient control of the infection in the future.

Clinical signs of BT appear as soon as five days post-infection. Therefore, in the early stages of an epidemic, infected animals are more quickly detected by clinical examination than by serology. In Italy, during the 2000-2001 BT-outbreak, sero-surveillance was only used in the decreasing phase of the epidemic curve (Giovannini et al., 2004). In a reporting system such as the one implemented during the course of the outbreak in Belgium, a succession of events has to occur before a case is detected. Theoretically, the reporting of suspect cases initiates an examination of the susceptible population which is under owner and veterinary observation. This first relies on the assumption that the infection will produce clinical signs; hence, subclinical cases will remain unnoticed (Doherr et al., 2001). BTV has in the past been isolated in several countries without clinical disease being recognised (Gibbs et al., 1994; Mulhern, 1985). Based on the sparse data from whole-herd-sampling during the northern European epidemic, it has been shown that a high proportion of cattle within a herd could be PCR or seropositive, while not showing any BT-clinical signs. Moreover, owners and veterinarians in Belgium had never previously experienced this exotic disease; therefore clinical signs were unfamiliar to them (Elbers et al., 2007). Also, owners may have been reluctant to report cases for fear of consequent loss of trade. The winter screening revealed indeed a higher prevalence than demonstrated by the

reporting of clinical cases. Results demonstrated a high level of exposure to BTV in the dairy herds. They confirm that BTV spreads very quickly in such an immunologically naïve ruminant population. The first Italian epidemic of BT in 2000 in Sardinia demonstrated a rate of spread of 30 km per week and 80% of the island eventually became infected. Both in Sardinia and Sicily, serological surveillance detected virus circulation to be more wide spread than indicated through clinical surveillance (Calistri et al., 2004). Serological screening demonstrated BT animal-prevalence levels ranging from 3.2 to 61.1% in Albania following recent introduction of infection (Di Ventura, 2004). In the present study, the obtained Pearson correlation coefficient shows that the spatial distribution in the two datasets is very similar, in the sense that there is a positive and strong linear relationship between both estimates. However, there are large scale-differences in estimated prevalences. The Pearson correlation coefficient indeed ignores the scales of the two sets of results (Dohoo et al., 2003).

In theory, each individual within the target population, namely the Belgian ruminant population, should have had an equal chance of being selected for sampling. For logistic reasons, only dairy herds with on-farm delivery of dairy product were included in the sampling frame. Moreover, only animals older than 24 months were sampled. Subpopulation sampling presents an opportunity for selection bias which must be accounted for when intending to extrapolate the results to the target population. From the outbreak data, sheep herds appeared to represent 57% of the total number of case herds. Ovine BTV infection cases might have been easier to detect since this species is commonly known to be more prone to develop the clinical form of the disease (Gibbs et al., 1994). However, the particularity of this BTV-8 epidemic seems to have been its ability to induce severe clinical signs in cattle too. Moreover, the analysis of the confirmation data showed that clinical signs observed in cattle were more specific than those observed in sheep (Toussaint et al., 2007b). In general, prevalence is known to be higher in cattle than in small ruminant populations (Ward et al., 1994; Di Ventura et al., 2004). On the other hand, a study conducted on the Indian sub-continent, demonstrated a higher prevalence in sheep than in the cattle population, with 45.7% and 33.4%, respectively (Sreenivasulu et al., 2004). Those findings demonstrate differences which can occur when sampling a particular species instead of another. In the same way, many studies have concluded older cattle are more likely to be positive to BTV antibodies than younger cattle, as a result of greater opportunity for repeated exposure to the virus (Uhaa et al., 1990; Ward et al., 1994; Lundervold et al., 2003). Factors such as breed-specific genetics or management methods differ a lot between beef and dairy cattle herds; hence, the level of prevalence may not follow identical patterns.

These findings currently provide the best information available on the unprecedented occurrence of BT in Belgium and emphasize the rapid and non-restricted spread of the virus in a susceptible ruminant population. Local variations in estimated prevalence should be further investigated to help identify particular risk factors and be able to better control future outbreaks. This study showed that the case detection system based on clinical suspicion underestimated the real impact of the epidemic, but provided an accurate indication of the spatial distribution of the infection.

ACKNOWLEDGEMENTS

This study is part of a contract research programme funded by the Belgian Federal Public Service of Health, Food Chain Safety, and Environment, the Belgian Fund for Health and Quality of Animals and Derived Products, and The Belgian Federal Agency for the Safety of the Food Chain. Partners in this programme are Avia-Gis (Zoersel, Belgium), the Gembloux

Agricultural University (Gembloux, Belgium), the Institute for Tropical Medicine (Antwerp, Belgium), Liège University (Liège, Belgium), the Veterinary and Agrochemical Research Centre (Brussels, Belgium), and the Walloon Agricultural Research Centre (Gembloux, Belgium). The authors gratefully acknowledge all the persons which have contributed to the collection of data for their cooperation in this study.

REFERENCES

Calistri, P., Giovannini, A., Conte, A., Nannini, D., Santucci, U., Patta, C., Rolesu, S. and Caporale, V.(2004). Bluetongue in Italy: Part I. Veterinaria Italiana 40, 243-251

Cannon, R.M., Roe, R.T. (1982). Livestock Disease Surveys: A field manual for veterinarians. Australian Bureau of Animal Health, Canberra

Caporale, V., MacLachlan, N.J., Pearson, J.E. and Schudel, A. (2004). Introduction Third International Symposium on Bluetongue. Veterinaria Italiana 40, 29-30

CRL, European Commission Reference Laboratory (2006). Bluetongue virus in the Netherlands identified as serotype 8 by Institute for Animal Health, Promed, August 28. Accessed at www.promedmail.org, archive no.: 20060828.2448

Daniels, P.W., Sendow, I., Pritchard, L.I., Sukarsih and Eaton, B.T. (2004). Regional overview of bluetongue viruses in South-East Asia: viruses, vectors and surveillance. Veterinaria Italiana 40, 94-100

Di Ventura, M., Tittarelli, M., Semproni, G., Bonfini, B., Savini, G., Conte, A. and Lika, A. (2004). Serological surveillance of bluetongue virus in cattle, sheep and goats in Albania. Veterinaria Italiana 40, 101-104

Doherr, M.G. and Audigé, L. (2001). Monitoring and surveillance for rare health-related events: a review from the veterinary perspective. Philosophical Transactions of the Royal Society B 356, 1097-1106

Dohoo, I., Martin, S.W. and Stryhn, H. (2003).Veterinary Epidemiologic Research. AVC Inc., Charlottetown

Dulac, G.C., Dubuc, C., Myers, D.J., Afshar, A. and Taylor, E.A. (1989). Incursion of Bluetongue virus type 11 and Epizootic Haemorrhagic disease of deer type 2 for 2 consecutive years in the Okanagan valley. Canadian Veterinary Journal 30, 351

EFSA (2007). Report on Epidemiological analysis of the 2006 bluetongue virus serotype 8 epidemic in north-western Europe: provisional findings through 31 January 2007. European Food Safety Authority. Accessed at:http://www.efsa.europa.eu/en/in_focus/bluetongue/-Bluetongue_report.s8.html

Elbers, A.R.W., Backx, A., Méroc, E., Gerbier, G., Staubach, C., Hendrickx, G., van der Speck, A. and Mintiens, K. (2007). Nature and severity of disease in sheep and cattle herds during the Bluetongue virus serotype 8 epidemic in 2006 in north-western Europe. Preventive Veterinary Medicine (in press)

FAO, Animal Production and Health Division (2006). Animal Health, disease cards, Bluetongue. Accessed at: http://www.fao.org/ag/againfo/subjects/fr/health/diseases-cards/bluetongue.html

Gerbier, G., Baldet, T., Tran, A., Hendrickx, G., Guis, H., Mintiens K., Ellbers, A. and Staubach, C. (2007). Modelling local dispersal of bluetongue serotype 8 using Random walk. Preventive Veterinary Medicine (in press)

Gerdes, G.H. (2004). A South African overview of the virus, vectors, surveillance and unique features of bluetongue. Veterinaria Italiana 40, 39-42

Gibbs, E.P.J. and Greiner, E.C. (1994). The epidemiology of bluetongue. Comparative Immunology, Microbiology and Infectious Diseases 17, 207-220

Hassan, A. (1992). Epidemiology of bluetongue virus infection in Malaysia. In: Walton, T.E., Osburn, B.I. (Eds.), Bluetongue, African horse sickness & Related Orbiviruses. CRC Press, Boca Raton, pp. 155-161

Herniman, K.A., Gumm, I.D., Owen, L., Taylor, W.P. and Sellers, R.F. (1980). Distribution of Bluetongue viruses and antibodies in some countries of the eastern hemisphere. Bulletin Office of International Epizootics 92, 581-586

MacLachlan, N.J. (2004). Bluetongue: pathogenesis and duration of viraemia. Veterinaria Italiana 40 (4), 462-467

Mo, C.L., Thompson, L.H., Homan, E.J., Oviedo, M.T., Greiner, E.C., Gonzales, J. and Saenz, M.R. (1994). Bluetongue virus isolations from vectors and ruminants in Central America and the Caribbean. American Journal of Veterinary Research 55, 211-215

Mulhern, F.J. (1985). Economic impact of bluetongue and related orbiviruses: Western Hemisphere. Progress in Clinical and Biological Research. 178, 21-25

OIE (2004). Manuel of Standards for Diagnostic Tests and Vaccines. Office Internationale des Epizooties 2004 Chapitre 2.1.9. Bluetongue

OIE Animal Health Department (2006). Bluetongue - Netherlands, Belgium, Germany-OIE. ProMed. August 21, 2006. Accessed at http://www.promedmail.org, archive no.: 20060821.2353

Qin, Q., Tai, Z., Wang, L., Luo, Z., Hu, J. and Lin, H. (1996). Bluetongue epidemiological survey and virus isolation in Xinjiang, China. In: St. George, T.D., Kegao, P. (Eds.), Bluetongue disease in Southeast Asia and Pacific, 1996, Proceedings of the First South-East Asia and Pacific Regional Bluetongue Symposium., ACIAR Proceedings (Australia) 66, 67-71

Rogan, W.J., Gladen, B. (1978). Estimating prevalence from the results of a screening test. American Journal of Epidemiology 107, 71-76

Sreenivasulu, D., Subba Rao, M.V., Reddy, Y.N. and Gard, G.P. (2004). Overview of bluetongue disease, viruses, vectors, surveillance and unique features: the Indian sub-continent and adjacent regions. VeterinariaItaliana 40, 73-77

Toussaint, J.F., Sailleau, C., Bréard, E., Zientara, S. and De Clercq, K. (2007). Bluetongue virus detection by two real-time RT-qPCRs targeting two different genomic segments. Journal Virology Methods 140, 115-123

Toussaint, J.F., Vandenbussche, F., Mast, J., Demeestere, L., Goris, N., Van Dessel, W., Vanopdenbosch, E., Kerkhofs, P., Zientara, S., Sailleau, C., Czaplicki, G., Depoorter, G., Dochy, J.M. and De Clercq, K. (2006). Bluetongue in Northern Europe. Veterinary Record 159, 327

Uhaa, I.J., Riemann, H.P., Thurmond, M.C. and Franti, C.E. (1990). A seroepidemiological study on bluetongue virus in dairy cattle in the central valley of California. Veterinary Research Communications 14, 99-112

Vandenbussche, F., Vanbinst, T., Verheyden, B., Van Dessel, W., Demeestere, L., De Poorter, G., Houdart, P., Bertels, G., Praet, N., Berkvens, D., Mintiens, K., Goris, N. and De Clerq, K. (2007). Evaluation of antibody-ELISA and real time RT-PCR for the diagnosis and profiling of bluetongue virus serotype 8 during the epidemic in Belgium in 2006. Veterinary Microbiology (in press)

Ward, M.P., Carpenter, T.E. and Osburn, B.I. (1994). Host factors affecting seroprevalence of bluetongue virus infections of cattle. American Journal of Veterinary Research 55, 916-920

FOOD SAFETY

A MULTISCALE MODEL OF E. COLI O157 TRANSMISSION IN THE SCOTTISH CATTLE POPULATION

L. MATTHEWS*, J. C. LOW, D. J. MELLOR, M. C. PEARCE, G. T. INNOCENT, G. J. GUNN AND S. W. J. REID

SUMMARY

Understanding the mechanisms by which zoonotic pathogens persist in animal reservoirs is key to the design of effective interventions. Enterobacteriaceae, which include *E. coli* O157, have the capacity to reside both in cattle and in environmental, wildlife and other livestock reservoirs. The within-herd dynamics of *E. coli* O157 infection have been subject to detailed analysis, but the mechanisms by which the pathogen persists in the national herd are not fully understood. By developing a model that captures within-herd transmission dynamics, herd-to-herd movement of infected animals and reservoirs of infection, we can quantify the theoretical contributions of different sources of infection to persistence in the national herd. By defining a threshold parameter for the persistence of infection in a cattle metapopulation, it is demonstrated that the combination of cattle movements and within-group transmission of infection cannot sustain *E. coli* O157 in the national herd in the absence of a reservoir of infection or high on-farm group-to-group transmission rates. Our results show that for scenarios with high intergroup transmission rates or reservoirs of infection that are dependent on prevalences of infection in the cattle population, observed dynamics are close to threshold behaviour. In this case, equilibrium prevalences in the national herd are sensitive to key parameters – herd-to-herd movement rates, group size and the within-group reproduction ratio, R_0. Combining transmission processes at multiple spatial scales – within-group, within-farm and farm-to-farm - provides a quantitative framework that can capture the impact of changes in dynamics at the within- or between-group scale on equilibrium prevalences of infection at a national scale. Our model therefore provides a coherent framework for the comparative analysis of alternative control measures that might be targeted at direct cattle-to-cattle, group-to-group or farm-to-farm transmission of infection.

INTRODUCTION

Escherichia coli O157 is a zoonotic pathogen capable of causing serious illness and mortality, and is a major cause of HUS (haemolytic uraemic syndrome) in children. *E. coli* O157 emerged in the 1980s as a significant cause of food-borne illness, but outbreaks in humans have also been linked to direct contact with livestock and the farm environment (Locking et al., 2001; Strachan et al., 2006).

Recent analyses (Matthews et al., 2006a; Matthews et al., 2006b) have developed the idea that super-shedders may drive the within-herd transmission dynamics. The capacity for a

* Louise Matthews, Faculty of Veterinary Medicine, University of Glasgow, Bearsden Road, Glasgow, G61 1QH. Email: l.matthews@vet.gla.ac.uk

pathogen to spread within its host population is quantified in terms of the basic reproduction ratio, R_0. This is the average number of infected individuals generated by a single infected individual introduced into a naïve population: if it is greater than 1, then on average the infection will spread; if it less than 1, outbreaks will tend to decline (Anderson & May, 1991). Quantification of the transmission dynamics of *E. coli* O157 suggests that the presence of super-shedding infections could be responsible for maintaining the basic reproduction ratio above 1 (Matthews et al., 2006a; Matthews et al., 2006b).

A number of options for the control of *E. coli* O157 in cattle have been proposed. These include vaccination (Dean-Nystrom et al., 2002; Potter et al., 2004), bacteriophage (Kudva et al., 1999; O'Flynn et al., 2004) and probiotics (Zhao et al., 1998; Zhao et al., 2003). Super-shedding infections provide a target for control options, and the direct application of therapeutic agents at the recto-anal junction (Naylor et al., 2007) is being evaluated as a means of control through directly targeting shedding at high densities. Changes in management practices are also being considered as methods of reducing cattle-to-cattle spread both within and between groups, and recent studies have shown that the two most promising factors for achieving a reduction in prevalence are maintaining the animals on dry bedding and keeping animals within the same group (Ellis-Iversen et al., 2007; Ellis-Iversen et al., 2008).

Design and implementation of the most effective control strategies to reduce prevalence at a national scale would benefit from a quantitative understanding of the transmission dynamics both within and between livestock holdings. Current models of the transmission of *E. coli* O157 have focused attention on the dynamics at a single scale: on within-host dynamics (Wood et al., 2006a; Wood et al., 2006b; Wood et al., 2006c); within-group dynamics (Matthews et al., 2006a; Matthews et al., 2006b; Wood et al., 2007) or herd-to-herd transmission of infection via cattle movements (Liu et al., 2007). However, simulation studies of pathogen spread within metapopulations have demonstrated the importance of combining an understanding of the within-group dynamics with the group-to-group movement of infected individuals (Park et al., 2001; Park et al., 2002; Cross et al., 2005; Cross et al., 2007). In such structured populations, critical values of R_0 required to allow disease invasion will exceed 1. Thus, a full understanding of the impact of control measures on the prevalence at a national scale requires a metapopulation model that combines both herd-to-herd movements and within-herd transmission dynamics.

In this paper, a quantitative framework that combines the key features of within-herd transmission dynamics, a reservoir of infection and herd-to-herd animal movements is developed. This framework permits the calculation of (i) a threshold for persistence and (ii) equilibrium prevalences of *E. coli* O157 infection in the Scottish national herd. The framework will be used to evaluate (i) the impact of alternative forms of reservoir on persistence in the national herd, (ii) the role of intergroup-transmission on persistence, and (iii) the sensitivity of equilibrium prevalences to key parameters.

MATERIALS AND METHODS

In this section, the transmission dynamics model, which captures both within-group transmission of *E. coli* O157 and animal movements to and from other livestock holdings, is outlined. This model framework is used to capture on-farm dynamics in terms of three characteristics: the expected prevalence when infected, p_{pos}; the expected period of infection in the group, T_{inf}; and the probability of the group being uninfected, p_0. Combined with estimates for cattle movement rates these quantities will be used to define a threshold parameter, R_{pop},

which must exceed 1 for infection to persist in the metapopulation of farms. Finally, an expression for equilibrium prevalences in the national herd is obtained.

Within-herd transmission dynamics

Cattle-to-cattle infection dynamics: The within-herd transmission dynamics are captured using a stochastic *SIS* (susceptible-infected-susceptible) description. New infections arise in the susceptible population via two routes: first, direct transmission from other infected individuals (with transmission rate, $\lambda S/N$, where λ is the force of infection and N the group size); and second, from a reservoir (at rate eS). Infected individuals are assumed to recover to the susceptible state at rate σI.

The model allows for host-to-host heterogeneity in infectiousness. The force-of-infection acting on susceptible individuals is therefore given by the transmission rate weighted by the relative infectiousness and summed over all infected individuals. Denoting the transmission rate by β, and the relative infectiousness of individual j, in a group of J infected individuals, by ω_j, the force of infection, λ, is therefore given by

$$\lambda = \sum_{j=1}^{J} \beta \omega_j \tag{1}$$

Host-to-host heterogeneity and super-shedding of *E. coli* O157: Previous analyses have demonstrated the potential role of super-shedders in the within-herd transmission dynamics. Fitting models to the distribution of prevalences demonstrated that, in the absence of other heterogeneities, 20% of infections could be responsible for 80% of the transmission (Matthews et al., 2006a). In this paper, persistence thresholds in the national herd are examined for both uniform and heterogeneous distributions of host infectiousness.

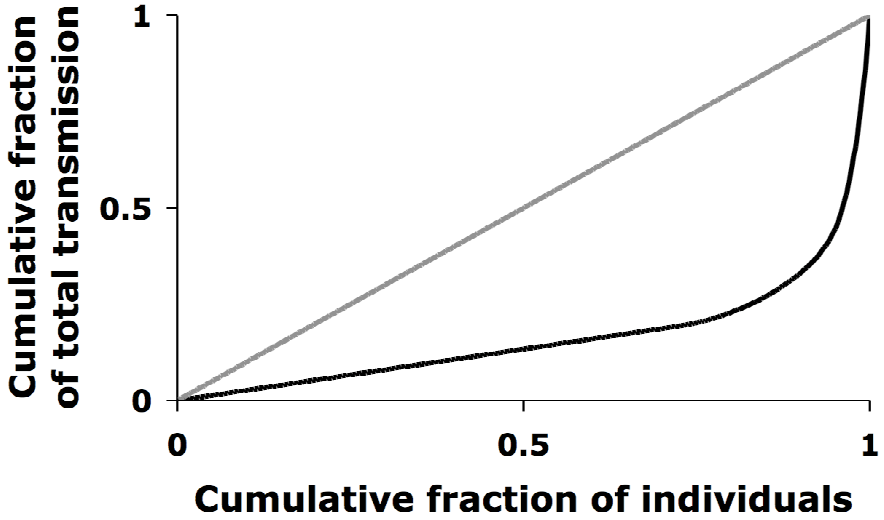


Fig. 1 Host-to-host variation in infectiousness for uniform infectiousness model (grey line) and heterogeneous infectious model (black line)

Figure 1 shows the relative contribution to the total transmission arising from host-to-host variation in infectiousness, based on our previous analyses of the prevalence data (Matthews et al., 2006b). In the case that all animals are assumed to be equally infectious, the cumulative contribution to the overall transmission increases linearly as the fraction of individuals included increases (grey line). In the case that animals are allowed to have differing degrees of infectiousness based on the observed distribution of bacterial loads, the best-fit model shows a non-linear increase in contribution to total transmission as the fraction of individuals included, ordered from low to high infectiousness, increases (black line).

Movement of animals into and out of herds: Movements-on of animals bring either susceptible or infected individuals to the population at the respective rates $k_{in}\ p_S N$ and $k_{in}\ p_I N$ where p_I and p_S are the proportions of infected and susceptible individuals in the national herd and k_{in} is the *per capita* cattle on-movement rate. Movements-off of animals result in susceptible and infected individuals leaving the herd at respective rates $k_{out}S$ and $k_{out}I$.

Cattle movement rates: Cattle movement rates were obtained from published estimates (Mitchell et al., 2005). The data summaries presented for lifetime cattle movements for 2002 and 2003 show that the mean number of movements per lifetime is approximately 2.1. Excluding movements to slaughter this gives a mean lifetime movement rate between farms of 1.1. Assuming a mean lifespan of 36 months, mean *per capita* movement rates (relative to a timescale based on the recovery period for *E. coli O157* infection of 3-4 weeks) in the range 0.02-0.03 are obtained. In this paper, an estimate of the mean movement rate of 0.025 (relative to the timescale for recovery) is used.

Demographic turnover: Susceptible and infected individuals are assumed to leave the herd as a consequence of natural mortality at respective rates μS and μI. Susceptible individuals are born at a rate bN. To maintain individual herd size it is assumed that $k_{in}+b=k_{out}+\mu$, and that this turnover rate has a fixed value for all livestock holdings.

Herd size distribution: Mean herd sizes for the Scottish cattle population were taken from the RADAR cattle book for 2006 (RADAR, 2006). Management group sizes were taken from a survey of Scottish beef finishing cattle conducted between 2002 and 2004 (Matthews et al., 2006a). Livestock holdings may have several management groups, but our default assumption is that these are managed separately. This basic model is later extended to explore the impact of inter-group transmission on the persistence of infection.

National herd size: Previous analyses have shown cattle movement-on and movement-off rates to be uncorrelated (Woolhouse et al., 2005). This assumption is made here, and, additionally, it is assumed that movement rates are uncorrelated with herd size. Maintenance of the national herd size in the model requires

$$\sum_i (k_{in})_i N_i = \sum_i (k_{out})_i N_i \tag{2}$$

The assumption of zero covariance between the movement rates and herd size gives

$$E((k_{in})_i N_i) = E((k_{in})_i)E(N_i) = k\bar{N} = E((k_{out})_i N_i) \tag{3}$$

where k and $\bar{N}$ are the mean movement rates and herd size, and thus the condition for maintenance of the national herd size is satisfied.

Within-herd dynamics: The differential equation model, which combines the within-herd dynamics and within-herd demographic turnover, takes the following form

$$\begin{aligned}\frac{dS}{dt} &= k_{in} p_S N - \mu S + bN - k_{out} S - \lambda S / N - eS + \sigma I \\ \frac{dI}{dt} &= k_{in} p_I N - \mu I - k_{out} I + \lambda S / N + eS - \sigma I\end{aligned} \tag{4}$$

As $S+I=N$, only one of the equations is required and can be simplified, using $k_{in}+b=k_{out}+\mu$, to the following

$$\frac{dS}{dt} = -k_{in} p_I S + (\sigma + k_{in} + b - k_{in} p_I) I - \lambda S / N - eS \tag{5}$$

This can be further simplified by defining the arrival rate, ρ, of infection in the herd by $\rho=k_{in}\, p_I$ and the effective recovery rate, σ', by $\sigma' = \sigma + k_{in} + b - \rho$, to the following

$$\frac{dS}{dt} = -\rho S + \sigma' I - \lambda S / N - eS \tag{6}$$

Stochastic simulation of within-herd dynamics: The stochastic model of within-herd dynamics is based on this deterministic framework (see Table 1 for the event types and their respective rates of occurrence). Taking this approach means that group-size, N, is preserved, since events are combined such that a movement-on, movement off, infection or recovery event is represented as the replacement of a susceptible by an infected animal and vice versa. This gives the following set of transitions

Table 1. Transitions for the stochastic within-herd SIS model

Transition	Symbolic notation	Rate
Susceptible to Infected	(S, I) to (S-1, I+1)	$\rho S + \lambda S/N + eS$
Infected to Susceptible	(S, I) to (S+1, I-1)	$\sigma' I$

Reservoirs of infection: A number of different formulations for the reservoir are explored: (i) a permanent reservoir of infection with a constant magnitude independent of infection prevalences in the cattle population, (ii) a permanent reservoir of infection with a magnitude proportional to mean prevalences in the national herd; and (iii) a transient reservoir dependent on local (on-farm) prevalences of infection.

Equilibrium distribution of within-herd prevalences: The parameter, ρ, the arrival rate of infection into a herd, where $\rho=k_{in}\, p_I$ is expected to be an important determinant of the properties of the dynamics within any given herd. The analysis therefore focused on the role of this parameter. For a given arrival rate of infection, ρ, the stochastic within-herd dynamics lead to an equilibrium distribution of the number of infected individuals. In the case that all individuals are equally infectious (ie $\lambda=\beta I$), the equilibrium distribution can be found analytically by iteratively solving

$$p_{J+1} = \frac{(\rho + e + \beta\frac{J}{N})(N-J)p_J}{\sigma'(J+1)} \tag{7}$$

for the probability, p_J, of observing J infected animals in the herd at the time of sampling. Fitting the heterogeneous version of the resulting probability distribution to a cross-sectional distribution of prevalences and obtaining estimates for β and ρ+e formed the basis of two previous analyses of the transmission dynamics of *E. coli* O157 (Matthews et al., 2006a; Matthews et al., 2006b).

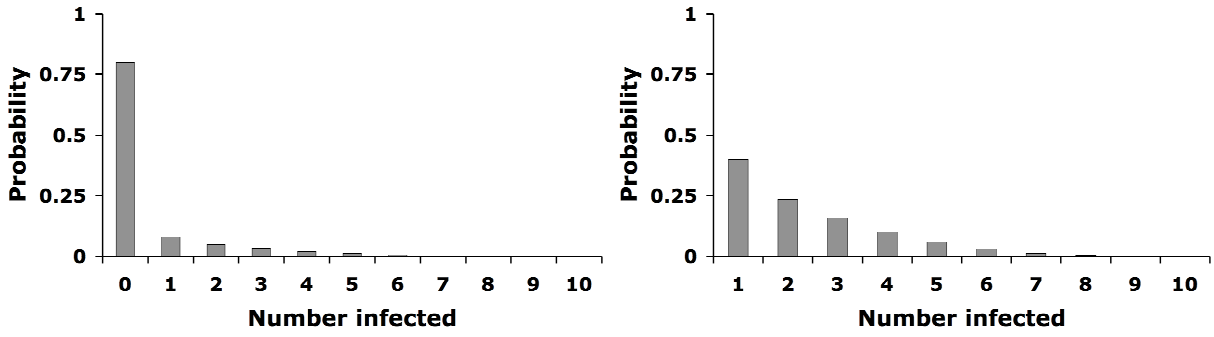


Fig. 2 Example probability distribution for the number of infected individuals for an SIS model in a group of 10 animals. Full distribution (left); conditional on infection being present (right)

An example of equilibrium distribution of states for an *SIS* model is shown in Fig. 2 (left) and the distribution of states conditional on the group having at least one infected individual is shown in Fig. 2 (right). The expected prevalences are denoted $p_{all}(\rho)$ and $p_{pos}(\rho)$ respectively. These are given by

$$p_{all}(\rho) = \sum_{J=0}^{N} p_J J \tag{8}$$

$$p_{pos}(\rho) = \frac{p_{all}}{(1-p_0)} \tag{9}$$

where $p_0(\rho)$ is the probability of the group being uninfected.

Within-herd persistence of infection: An additional property of the within-herd equilibrium state is now defined. This is $T_{inf,}$ the mean time the herd remains infected following introduction of infection. The quantity T_{inf} and the quantity $1/\rho N$, the mean waiting time for the arrival of a new infection, must satisfy

$$\frac{T_{\inf}}{1/\rho N} = \frac{1-p_0}{p_0} \tag{10}$$

Herd contribution: Using these characteristics, the contribution of a herd, $R(\rho)$, is defined by the product of its expected prevalence when infected, the herd size, the rate of movement of animals out of the herd, and the expected time the herd is infected

$$R(\rho) = T_{\inf}(\rho) p_{pos}(\rho) k_{out} N \tag{11}$$

This quantity gives the expected number of infected animals leaving the group during the period the group remains infected, following the introduction of infection into a fully susceptible group. It is shown below that this quantity forms the basis for a threshold for persistence of infection in the farm metapopulation.

Metapopulation transmission dynamics

Equilibrium prevalence in a metapopulation of herds: To develop a threshold for persistence, consider initially a population of identical herds, of size N, each undergoing within-herd transmission dynamics and moving animals in and out of the herds as specified in the above model description. At equilibrium, the expected prevalence of infection in animals leaving herds will equal the expected prevalence entering herds

$$\rho N = p_{all}(\rho) k_{out} N \equiv (1 - p_0(\rho)) p_{pos}(\rho) k_{out} N \tag{12}$$

Rearranging Eq.(10) to isolate ρN and substituting into Eq.(12) gives

$$\frac{1}{p_0(\rho)} = T_{\inf}(\rho) p_{pos}(\rho) k_{out} N \tag{13}$$

or, using Eq.(11)

$$R(\rho) = \frac{1}{p_0(\rho)} \tag{14}$$

Thus, the equilibrium prevalence in the national herd, p_I, (which determines the value of ρ via $\rho = p_I k_{in}$) must satisfy this expression[1].

Threshold for persistence of *E. coli* O157 in the national herd: Equation (14) is now used to consider thresholds for persistence in the metapopulation. As $p_0(\rho)$ is the probability of a herd being uninfected, whenever infection is present is the population of herds $p_0(\rho)$ must be less than 1. Thus, $1/p_0(\rho)$ and therefore $R(\rho)$ will be greater than 1 whenever infection is present. The critical value determining whether infection can persist in the population is given by the value of $R(\rho)$ as p_I and hence ρ tend to zero, which is denoted R_{pop}.

If $R_{pop} = R(0) = \lim_{\rho \to 0} R(\rho) > 1$ infection can persist in the population of herds.

[1] Note the analogy to the equilibrium state for deterministic *SIS* dynamics in a single herd: at equilibrium $R_0 = 1/x$ where $x = S/N$ is the proportion of animals that are uninfected.

If $R_{pop} = R(0) = \lim_{\rho \to 0} R(\rho) < 1$ infection cannot persist in the population of herds.

The quantity R_{pop} can be viewed as the metapopulation version of the within-herd basic reproduction ratio, R_0.

Heterogeneous populations of herds: These results can be readily generalised to accommodate a heterogeneous population. Generalising the invasion threshold to account for heterogeneities between farms requires the transmissibility, $R_i(0)$, of holding i to be weighted by its relative 'susceptibility' $(k_{in})_i N_i / E(k_{in} N)$ to the arrival of infection. Denoting the expectation of X by $E(X)$, this generates

$$R_{pop} = \frac{E(k_{in} N R(0))}{E(k_{in} N)} \tag{15}$$

or, adopting the notation of network analysis

$$R_{pop} = \frac{\langle k_{in} N R(0) \rangle}{\langle k_{in} N \rangle} \tag{16}$$

where $R(0)$, k_{in} and N are livestock holding specific. Note that when susceptibilities and transmissibilities are attributed solely to on and off movement rates, this is equal to the standard percolation threshold on a random network

$$R_{pop} = \frac{\langle k_{in} k_{out} \rangle}{\langle k_{in} \rangle} \tag{17}$$

It is straightforward to show that Eq.(16) for R_{pop} has the same threshold properties that were demonstrated for the uniform population, namely

If $R_{pop} = \frac{\langle k_{in} N R(0) \rangle}{\langle k_{in} N \rangle} > 1$ infection can persist in the population of herds.

If $R_{pop} = \frac{\langle k_{in} N R(0) \rangle}{\langle k_{in} N \rangle} < 1$ infection cannot persist in the population of herds.

Exploiting the lack of covariance between k_{in}, N and k_{out}, Eq.(16) simplifies to

$$R_{pop} = \frac{\langle N R(0) \rangle}{\langle N \rangle} \tag{18}$$

Approximations for low movement rates: For low movement rates and low prevalences, mean prevalences on any given farm are expected to depend linearly on the movement-on rate, k_{in}. Exploratory simulations also demonstrate that the expected on-farm prevalence when infected, p_{pos}, is relatively independent of k_{out}. Therefore, the number of infected individuals

leaving a group during its period of infection depends linearly on the movement-off rate. These approximately linear dependencies and our assumption of zero covariance between k_{in} and k_{out} means that our calculations of equilibrium prevalences can be simplified by using mean values of the movement rates.

Parameter estimates: The results of previous analyses of cross-sectional prevalence data were used to parameterise the model (Matthews et al., 2006a; Matthews et al., 2006b). These analyses used the stochastic model described in the within-herd dynamics section to generate equilibrium distributions of prevalence. These were compared with observed prevalence distributions and maximum likelihood methods were used to obtain estimates of the transmission rate, β, (and hence R_0) and the immigration rate, which in the current notation is the sum of the reservoir and movement-on components ie $\rho+e$. The analysis of the metapopulation equilibrium dynamics conducted here is based on parameter estimates (see Table 2) obtained via this methodology for both the uniform and heterogeneous infectiousness (super-shedder) models (see Fig.1).

Table 2. Maximum likelihood parameter estimates

	Basic reproduction ratio, R_0	Immigration rate, ρ+e
Uniform infectiousness model	1.15	0.005
Heterogeneous infectiousness model	1.5	0.009

RESULTS

Transmission dynamics

Thresholds for persistence: The threshold parameter, R_{pop}, determines whether infection can be maintained in the metapopulation of cattle holdings. Analogously to the within-group threshold parameter, R_0: if $R_{pop}>1$, infection can persist in the metapopulation; if $R_{pop}<1$ infection cannot persist in the metapopulation.

Parameter estimates for the within-group basic reproduction number, R_0, taken from previous analyses of *E. coli* O157 transmission dynamics (see Table 2) were used to obtain estimates for the threshold parameter, R_{pop} in the absence of a reservoir of infection. Results are shown for both the super-shedder model (black lines) and uniform infectiousness models (grey lines). Figure 3a shows that R_{pop} is below 1 until very high movement rates are achieved (approximately 0.6 on a timescale relative to the recovery period from *E. coli* O157 infection). When heterogeneity in management group size is included (dotted lines), instead of using a mean management group size of 18 (solid lines), the threshold is shifted towards lower movement rates. For observed mean movement rates in the range 0.02-0.03 (relative to the timescale of recovery from infection), R_{pop} is substantially less than 1 for all four scenarios considered.

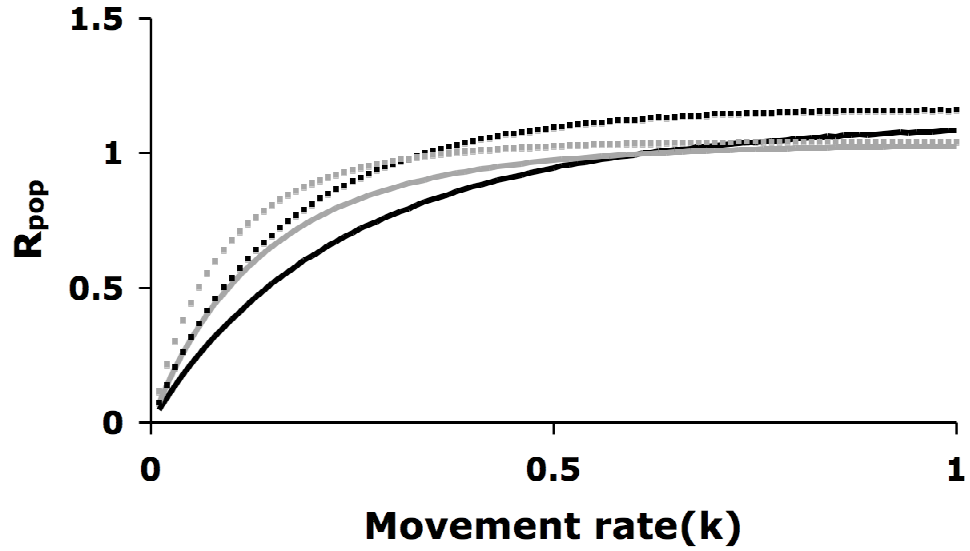


Fig. 3a Dependence of the persistence threshold, R_{pop}, on mean movement rate in the absence of reservoirs of infection for the uniform infectiousness model (uniform group size (grey line); heterogeneous group size distribution (dotted grey line)) and the super-shedder model (uniform group size (black line); heterogeneous group size distribution (dotted black line))

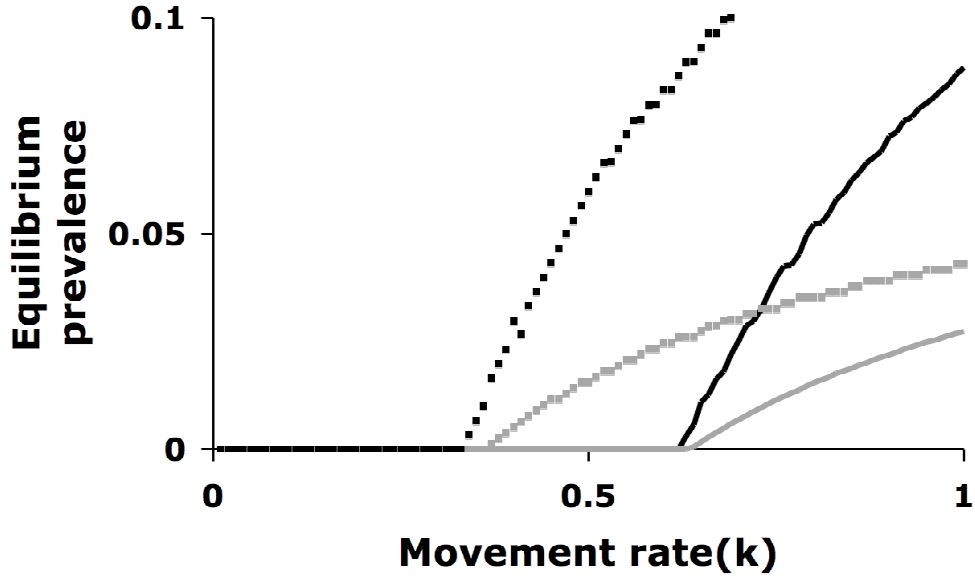


Fig. 3b Equilibrium prevalence versus mean movement rate for the uniform infectiousness model (uniform group size (grey line); heterogeneous group size distribution (dotted grey line)) and the super-shedder model (uniform group size (black line); heterogeneous group size distribution (dotted black line))

Equilibrium prevalence of infection: The role of the persistence parameter R_{pop} is reflected in the corresponding equilibrium prevalences of infection. Figure 3b (in conjunction with Fig. 3a) demonstrates that for values of the movement rate for which $R_{pop}<1$, the equilibrium prevalence is zero. Once R_{pop} exceeds 1 a sharp rise in prevalence of infection is observed, increasing to a plateau prevalence (not shown). In the absence of a reservoir, observed movement rates are insufficient to sustain infection in the national herd.

Persistent reservoirs of infection: Persistent (non-decaying) reservoirs of infection are incorporated in the model in two ways (i) a fixed reservoir, independent of the infection prevalence in the cattle population (Fig. 4a,b solid grey lines), and (ii) a reservoir with a magnitude proportional to the equilibrium prevalence in the national herd (Fig 4a,b dotted grey lines). Figure 4a (super-shedder model) and fig. 4b (uniform infectiousness model) demonstrate that both types of reservoir can generate observed national prevalences (of approximately 4% (Matthews et al., 2006a)) at low movement rates in the range consistent with observation.

The magnitude of the fixed reservoir was specified by our previous parameter estimates for the total immigration rate (see Table 2). The expected prevalence which would arise from this reservoir alone, without cattle to cattle transmission, is shown by the horizontal dashed grey lines. The expected prevalence arising from the combination of cattle-to-cattle transmission and herd-to-herd movements in the absence of a reservoir is shown by the black lines. At low movement rates, although cattle movements and cattle-to-cattle transmission are insufficient to maintain infection in the national herd, these processes act to amplify the infection levels arising from the reservoir (the solid grey line versus the horizontal dashed grey line).

Observed prevalences are much more sensitive to changes in movement rates for the proportional reservoir than the fixed reservoir, with equilibrium prevalences tending to zero as movement rates decline.

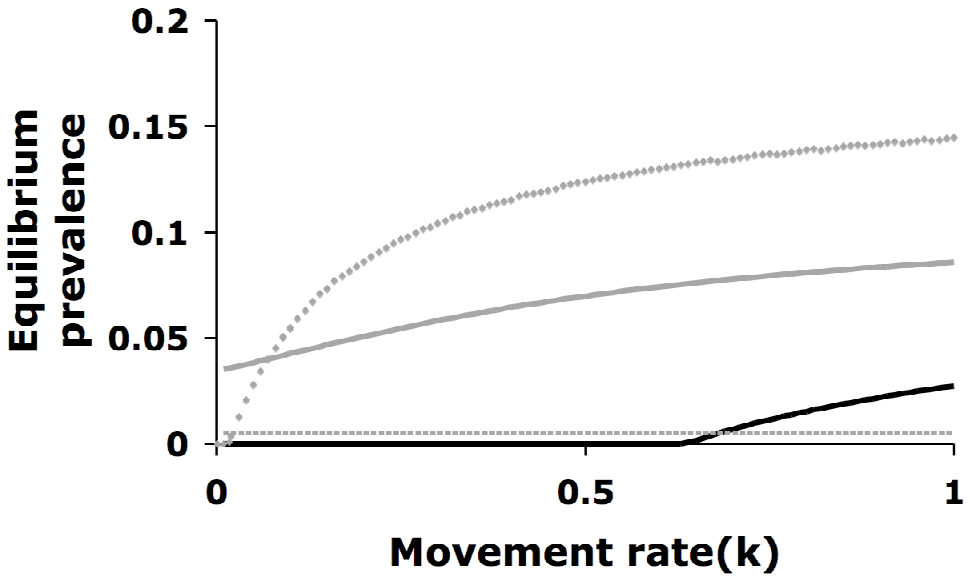


Fig. 4a Equilibrium prevalences in the national herd for the uniform infectiousness model under three different scenarios: no reservoir (black line); a global reservoir proportional to the equilibrium prevalence (dotted grey); and a constant background reservoir (solid grey line). The horizontal dashed grey line shows the expected prevalence from the constant background reservoir in the absence of cattle-to-cattle transmission

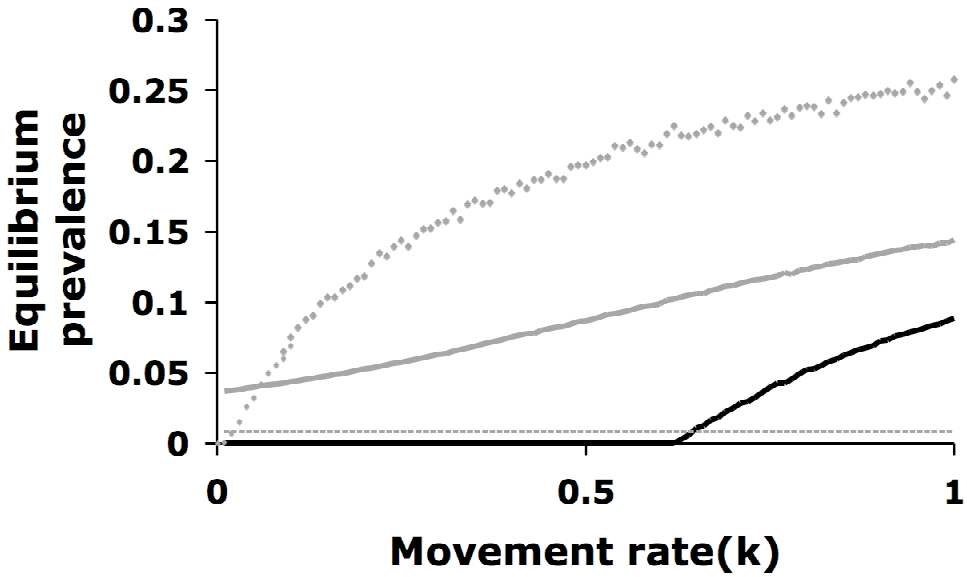


Fig. 4b Equilibrium prevalences in the national herd for the super-shedder model under three different scenarios: no reservoir (black line); a global reservoir proportional to the equilibrium prevalence (dotted grey); and a constant background reservoir (solid grey line). The horizontal dashed grey line shows the expected prevalence from the constant background reservoir in the absence of cattle-to-cattle transmission

Transient reservoirs: Scenarios were also considered in which reservoirs of infection depend not on national equilibrium prevalences but on local, on-farm prevalences which generate a local reservoir that decays over time. Selected decay rates are 0.02, 0.05, 0.1 and 0.2 (relative to the timescale for recovery from infection). Infectivity is assumed to build up in the environment at a rate proportional to current prevalences of infection. For each decay rate, the rate of accumulation of infectivity has been selected to generate observed prevalences (approximately 4%), at observed mean movement rates, when group size equals the mean observed management group size.

Figure 5a shows the equilibrium prevalence in the national herd, for the super-shedder model, as a function of group size, for the four reservoir decay rates. At a group size of 18 (the observed mean management group size), each scenario generates an equilibrium prevalence of approximately 4%, consistent with observation. There is a threshold group size below which infection cannot be sustained in the national herd, and substantial increases in prevalence as group size increases. Close to the threshold value of N, equilibrium prevalences depend sensitively on the group size.

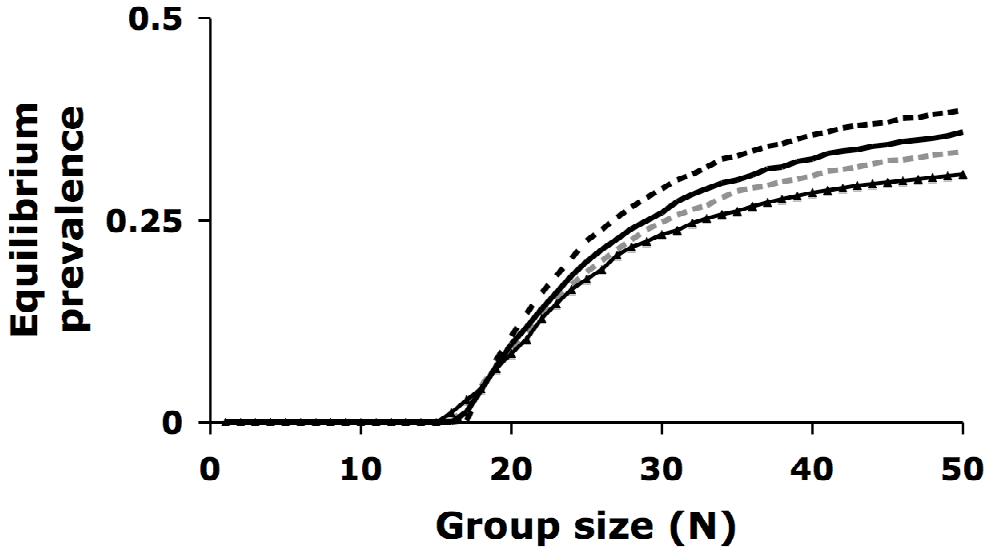


Fig. 5a Equilibrium prevalences in the national herd, for the super-shedder model, under the assumption of a locally acting on-farm reservoir of infection with decay rates of 0.02 (black plus triangles), 0.05 (dashed grey), 0.1 (black) and 0.2 (dashed black)

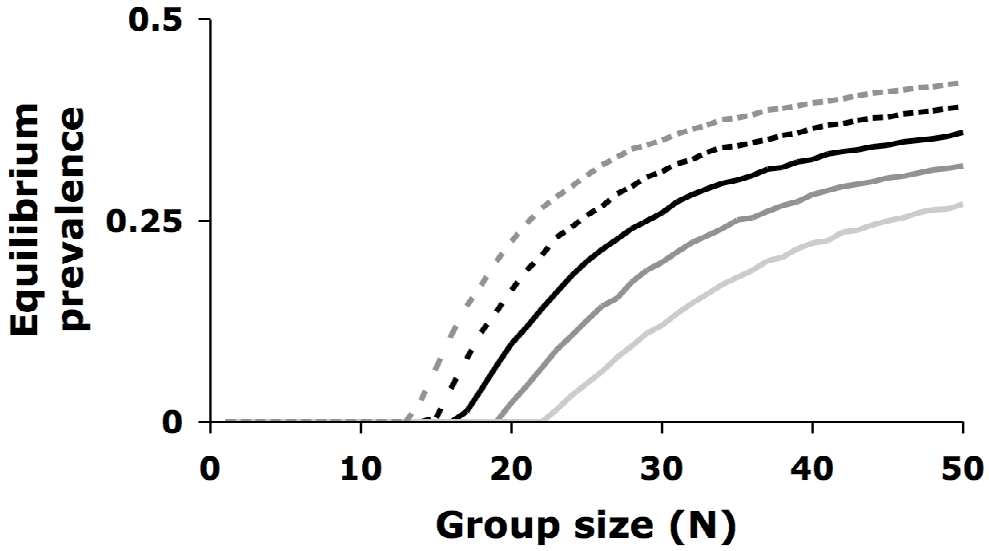


Fig. 5b Equilibrium prevalences in the national herd, for the super-shedder model, under the assumption of a locally acting on-farm reservoir of infection with decay rate=0.1. Curves are shown for the uniform infectiousness model with R_0 values of 1.7 (dashed grey), 1.6 (dashed black), 1.5 (solid black), 1.4 (dark grey) and 1.3 (light grey)

Figure 5b shows the dependence of these equilibrium prevalences on the within-herd reproduction ratio, R_0, selected in the range 1.3 - 1.7. At higher R_0 values, the threshold value of N required for persistence is reduced, and the plateau equilibrium prevalence is higher. The sensitivity of the equilibrium prevalence to changes in R_0 means that close to the persistence threshold, relatively small changes in R_0 could result in substantial changes in the national equilibrium prevalence..

Inter-group transmission: The scenarios described above assume that different cattle groups on a farm are managed separately such that inter-group transmission of infection does not occur. Scenarios are now considered in which a specified proportion of the transmission from any given infected animal is assumed to be to animals outside the group. Figure 6 shows the equilibrium prevalences as a function of group size for a herd comprising eight cattle groups (when N=18, this gives a realistic herd size of 144). The R_0 value is scaled upwards to account for transmission outside of the group, to give a net within-group R_0 equal to our estimated values (see Table 2). For observed mean group sizes of 18 animals, the observed prevalence of approximately 4% requires an inter-group transmission proportion of 0.21 or greater. Thus, high rates of inter-group transmission could be sufficient to sustain observed prevalences in the absence of other reservoirs of infection.

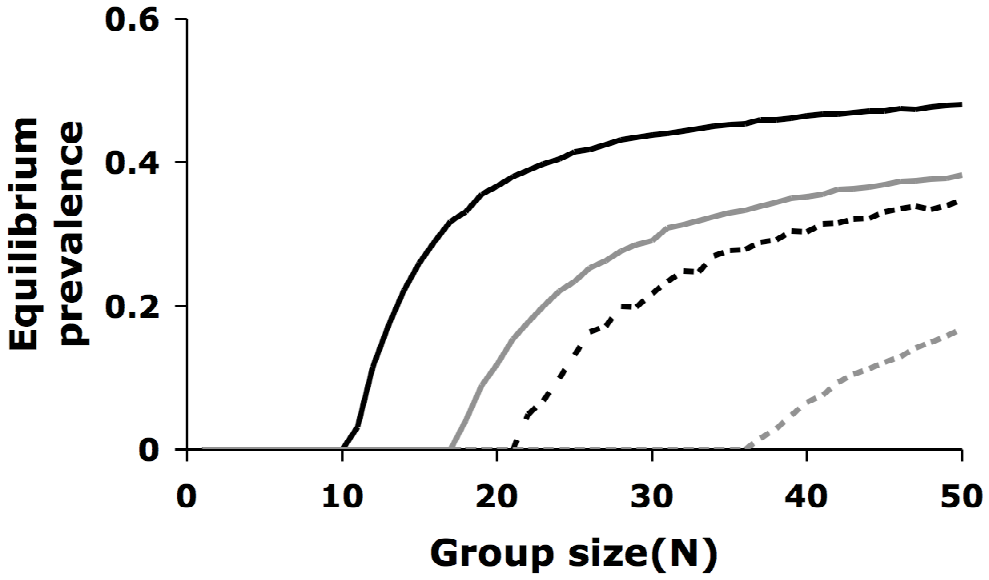


Fig. 6 Equilibrium prevalences in the national herd, for the super-shedder model, under the assumption of on-farm group-to-group transmission. Curves are shown for inter-group transmission proportions of 0.21 (grey), 0.3 (black), 0.15 (dashed black) and 0.1 (dashed grey) and net within-group R_0 values of 1.5

DISCUSSION

This paper describes the development of a metapopulation model of the transmission dynamics of *E. coli* O157 that combines dynamic processes at the within-group, within-herd and herd-to-herd scales. Specifically, by combining cattle movements, cattle-to-cattle transmission and reservoirs of infection, this framework allows the examination of potential mechanisms of persistence of *E. coli* O157 in the national herd, and the sensitivity of the national prevalence to changes in key parameters under alternative transmission scenarios to be evaluated.

An expression for a threshold parameter, R_{pop}, which determines whether infection can persist in the cattle metapopulation is developed. If R_{pop} exceeds 1 infection can persist; if R_{pop} is less than 1 infection will become extinct. By evaluating R_{pop} using parameter values estimated for the *E. coli* O157 transmission dynamics in the Scottish cattle population, it is demonstrated that in the absence of either a reservoir of infection or high inter-group transmission rates, R_{pop} is

less than 1. This suggests that observed cattle movement rates are too low to sustain infection in the national herd without additional sources of infection, which is consistent with the results of a preliminary exploration of the factors contributing to the metapopulation persistence of *E. coli* O157 (Liu et al., 2007). Potential alternative sources of infection include reservoirs of infection arising from the environment, wildlife or other livestock or group-to-group transmission of infection within the livestock holding, which would arise through close contact between animals in different groups or shared use of pasture.

It is demonstrated that introducing a reservoir allows infection to persist in the national herd at observed management group sizes and cattle movement rates. A number of reservoir scenarios, parameterised to generate prevalences consistent with observed values in the national herd, have been explored. A background reservoir of constant magnitude generates equilibrium prevalences which are least sensitive to changes in parameter values. This type of reservoir would occur if *E. coli* O157 infection could persist permanently in the environment or other livestock or wildlife host, in the absence of infection in the cattle population.

A permanent reservoir with a magnitude proportional to the prevalence of infection in the national herd was also considered. This type of reservoir would reflect a situation in which infection in cattle could generate infection in the environment or other wildlife or livestock hosts, but infection could not be maintained in those other reservoirs in the absence of infection in cattle. Such a reservoir would require additional input (for example, from a small background reservoir) to establish initially, but once established would be able to sustain infection in the national herd when $R_{pop}<1$.

The final reservoir scenario considered was a local, transient reservoir. In this case, shedding of bacteria by cattle is assumed to generate a temporary reservoir of infectivity in the environment. Reservoirs with persistence times ranging from five to fifty times the typical recovery timescale are shown to be able to generate prevalences consistent with observation.

As an alternative to a reservoir of infection, the role of inter-group transmission on the metapopulation persistence of infection was also evaluated. Our results show that for sufficiently high inter-group transmission rates persistence of infection at observed prevalences can be obtained without a reservoir of infection.

For each of the reservoir and inter-group transmission scenarios, the sensitivity of the equilibrium prevalences to several key parameters - the movement rate, the management group size and the within-group reproduction ratio R_0 – was examined. With the exception of the fixed background reservoir, these analyses show the dynamics to be close to threshold behaviour, meaning that relatively small changes in parameters can produce substantial changes in prevalence.

In summary, the model developed in this paper provides a framework that allows within-group, group-to-group and farm-to-farm dynamics to be combined. This multiscale approach provides a means of assessing the impact of changes in within-host carriage and host-to-host transmission on the expected prevalence of infection at a national scale. This model therefore provides a quantitative tool for assessing and comparing the efficacy of control measures targeted at different intervention points in the transmission dynamics.

ACKNOWLEDGEMENTS

The authors are grateful to the Wellcome Trust for financial support through an International Partnership Research Award in Veterinary Epidemiology (IPRAVE), *Epidemiology and Evolution of Enterobacteriaceae Infections in Humans and Domestic Animals* and a Mathematical Biology Research Training Fellowship (to LM).

REFERENCES

Anderson, R. M. and May, R. M. (1991) Infectious Diseases of Humans: Dynamics and Control, Oxford University Press, Oxford.

Cross, P. C., Johnson, P. L. F., Lloyd-Smith, J. O. and Getz, W. M. (2007) Utility of R-0 as a predictor of disease invasion in structured populations. J. Roy. Soc. Interface 4**,** 315-324.

Cross, P. C., Lloyd-Smith, J. O., Johnson, P. L. F. and Getz, W. M. (2005) Duelling timescales of host movement and disease recovery determine invasion of disease in structured populations. Ecology Letters 8**,** 587-595.

Dean-Nystrom, E. A., Gansheroff, L. J., Mills, M., Moon, H. W. and O'Brien, A. D. (2002) Vaccination of pregnant dams with Intimin(O157) protects suckling piglets from Escherichia coli O157 : H7 infection. Infect. Immun. 70**,** 2414-2418.

Ellis-Iversen, J., Smith, R. P., Snow, L. C., Watson, E., Millar, M. F., Pritchard, G. C., Sayers, A. R., Cook, A. J. C., Evans, S. J. and Paiba, G. A. (2007) Identification of management risk factors for VTEC O157 in young-stock in England and Wales. Prev. Vet. Med. 82**,** 29-41.

Ellis-Iversen, J., Smith, R. P., Van Winden, S., Paiba, G. A., Watson, E., Snow, L. C. and Cook, A. J. C. (2008) Farm practices to control E. coli O157 in young cattle - A randomised controlled trial. Vet. Res. 39. In press.

Kudva, I. T., Jelacic, S., Tarr, P. I., Youderian, P. and Hovde, C. J. (1999) Biocontrol of Escherichia coli O157 with O157-specific bacteriophages. Appl. Envir. Microbiol. 65**,** 3767-3773.

Liu, W. C., Matthews, L., Chase-Topping, M., Savill, N. J., Shaw, D. J. and Woolhouse, M. E. J. (2007) Metapopulation dynamics of Escherichia coli O157 in cattle: an exploratory model. J. Roy. Soc. Interface 4**,** 917-924.

Locking, M. E., O'Brien, S. J., Reilly, W. J., Wright, E. M., Campbell, D. M., Coia, J. E., Browning, L. M. and Ramsay, C. N. (2001) Risk factors for sporadic cases of Escherichia coli O157 infection: the importance of contact with animal excreta. Epidemiol. Infect. 127**,** 215-220.

Matthews, L., Low, J. C., Gally, D. L., Pearce, M. C., Mellor, D. J., Heesterbeek, J. A., Chase-Topping, M., Naylor, S. W., Shaw, D. J., Reid, S. W., Gunn, G. J. and Woolhouse, M. E. (2006a) Heterogeneous shedding of Escherichia coli O157 in cattle and its implications for control. Proc. Natl. Acad. Sci. U S A 103**,** 547-52.

Matthews, L., McKendrick, I. J., Ternent, H., Gunn, G. J., Synge, B. and Woolhouse, M. E. (2006b) Super-shedding cattle and the transmission dynamics of Escherichia coli O157. Epidemiol. Infect. 134**,** 131-42.

Mitchell, A., Bourn, D., Mawdsley, J., Wint, W., Clifton-Hadley, R. and Gilbert, M. (2005) Characteristics of cattle movements in Britain - an analysis of records from the Cattle Tracing System. Anim. Sci. 80**,** 265-273.

Naylor, S. W., Nart, P., Sales, J., Flockhart, A., Gally, D. L. and Low, J. C. (2007) Impact of the direct application of therapeutic agents to the terminal recta of experimentally colonized calves on Escherichia coli O157 : H7 shedding. Appl. Envir. Microbiol. 73**,** 1493-1500.

O'Flynn, G., Ross, R. P., Fitzgerald, G. F. and Coffey, A. (2004) Evaluation of a cocktail of three bacteriophages for biocontrol of Escherichia coli O157 : H7. Appl. Envir. Microbiol. 70**,** 3417-3424.

Park, A. W., Gubbins, S. and Gilligan, C. A. (2001) Invasion and persistence of plant parasites in a spatially structured host population. Oikos 94**,** 162-174.

Park, A. W., Gubbins, S. and Gilligan, C. A. (2002) Extinction times for closed epidemics: the effects of host spatial structure. Ecology Letters 5**,** 747-755.

Potter, A. A., Klashinsky, S., Li, Y. L., Frey, E., Townsend, H., Rogan, D., Erickson, G., Hinkley, S., Klopfenstein, T., Moxley, R. A., Smith, D. R. and Finlay, B. B. (2004) Decreased shedding of Escherichia coli O157 : H7 by cattle following vaccination with type III secreted proteins. Vaccine 22**,** 362-369.

RADAR (2006) www.defra.gov.uk/animalh/diseases/vetsurveillance/reports/population.htm.

Strachan, N. J. C., Dunn, G. M., Locking, M. E., Reid, T. M. S. and Ogden, I. D. (2006) Escherichia coli O157: Burger bug or environmental pathogen? Int. J. Food Microbiol. 112**,** 129-137.

Wood, J. C., McKendrick, I. J. and Gettinby, G. (2006a) A simulation model for the study of the within-animal infection dynamics of E-coli O157. Prev. Vet. Med. 74**,** 180-193.

Wood, J. C., McKendrick, I. J. and Gettinby, G. (2006b) Assessing the efficacy of within-animal control strategies against E-coli O157: A simulation study. Prev. Vet. Med. 74**,** 194-211.

Wood, J. C., McKendrick, I. J. and Gettinby, G. (2007) A simulation model to assess herd-level intervention strategies against E-coli O157. Epidemiol. Infect. 135**,** 749-764.

Wood, J. C., Speirs, D. C., Naylor, S. W., Gettinby, G. and McKendrick, L. J. (2006c) A continuum model of the within-animal population dynamics of E-coli O157. J. Biol. Sys. 14**,** 425-443.

Woolhouse, M. E. J., Shaw, D. J., Matthews, L., Liu, W. C., Mellor, D. J. and Thomas, M. R. (2005) Epidemiological implications of the contact network structure for cattle farms and the 20-80 rule. Biology Letters 1**,** 350-352.

Zhao, T., Doyle, M. P., Harmon, B. G., Brown, C. A., Mueller, P. O. E. and Parks, A. H. (1998) Reduction of carriage of enterohaemorrhagic Escherichia coli O157 : H7 in cattle by inoculation with probiotic bacteria. J. Clin. Microbiol. 36**,** 641-647.

Zhao, T., Tkalcic, S., Doyle, M. P., Harmon, B. G., Brown, C. A. and Zhao, P. (2003) Pathogenicity of enterohemorrhagic Escherichia coli in neonatal calves and evaluation of fecal shedding by treatment with probiotic Escherichia coli. J. Food Protect. 66**,** 924-930.

DETECTION OF ANTIMICROBIAL RESISTANT *SALMONELLA* IN LIVESTOCK: A CHANCE EVENT?

K. KAVANAGH*, L. KELLY, E.L. SNARY AND G. GETTINBY

SUMMARY

In Great Britain, monitoring the levels of antimicrobial resistant Salmonella in livestock occurs both as part of a passive surveillance system and as structured surveys. To provide insight into such surveillance activities, a probabilistic model has been developed to assess the probability of detecting resistance at the faecal, pen and farm level. Using this model, it is concluded that the probability of detecting resistant Salmonella is dependent upon the level of resistance within sample/pen/farm and the diagnostic power of the test used. The likelihood of detecting low level (e.g. emerging) resistance on individual farms was low and therefore the use of selective plating (antimicrobial present in the plate at the specified breakpoint concentration so growth confirms the presence of resistant Salmonella) is recommended. Importantly, the models provide an insight into the sampling and testing methods and could therefore be used to inform any future on-farm surveillance programmes or research projects.

INTRODUCTION

The extent to which farmed animals are infected with antimicrobial resistant zoonotic bacteria such as *Salmonella* is of concern due to the potential exposure of humans to an additional pool of resistance genes via the food chain (Swann Report, 1969). However, detection of such resistance may be difficult especially when the prevalence of resistant infection is low as with, for example, emerging resistant strains. Probabilistic methods can be used to model the likelihood of detecting them in faecal samples. This approach has been used for *Salmonella* in individual and pooled faecal samples (Arnold et al., 2005) and resistant organisms in general, for different levels of resistance in the host (Davison et al., 2000). In this paper, theoretical probabilistic models are described that could be used to assess the sensitivity of antimicrobial resistance testing in surveillance schemes or research projects. The probabilistic models are formulated to account for the method of faecal sample collection on-farm. By modelling the sampling system together with the variability within the testing process, the probability of detecting resistant *Salmonella* on an individual farm is assessed. This, alongside calculations of the minimum prevalence of resistance which the current surveillance system can detect, provides insight into on-farm surveillance activities (passive or active) for antimicrobial resistance, particularly if a low level of resistance is present.

MATERIALS AND METHODS

* Kim Kavanagh, Department of Statistics and Modelling Science, University of Strathclyde, 24 Richmond Street, Glasgow, G1 1XH, UK. Email: kim@stams.strath.ac.uk

The probability of detection is assessed in a hypothetical farm where animals are kept in penned groups; for illustrative purposes here a pig farm is considered. Faecal samples can be collected in a variety of ways, for example individual pigs sampled and these samples each tested for resistant Salmonella or multiple pigs sampled and the samples pooled into a composite sample and tested (see Fig. 1). The aims of these two types of sampling methods are different as the former provides information at the individual animal level and the latter at the pen or herd level. Given the prevalence of resistant infection within the pen, p and a proportion of resistant Salmonella organisms within the composite sample, pc, the probability of detecting resistance at the faecal sample level, the pen level and the farm level is formulated.

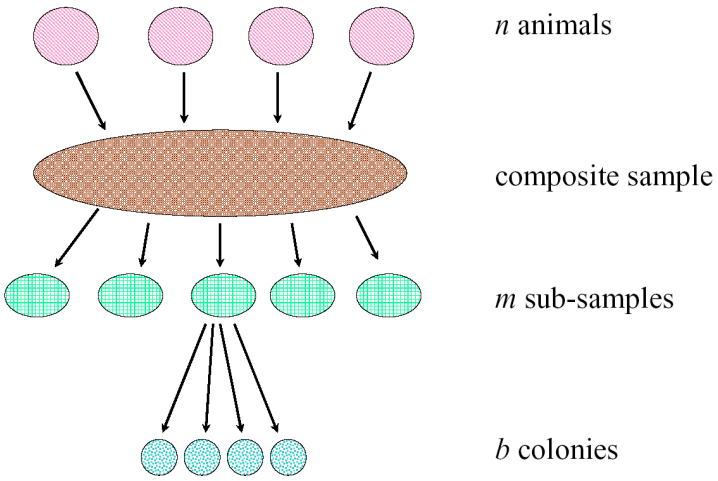


Fig. 1. Illustration of the sampling structure for a composite sample

Individual Sample Model

Assume that faecal samples from n pigs are collected from each pen and that each faecal sample is examined individually. It is assumed that an individual infected pig is infected with only a single resistant phenotype of Salmonella. Given this, the faeces produced by an individual contains Salmonella which are either all sensitive or all resistant to the antimicrobial in question. This implies that, if the test is perfect,

$$Pr(\text{detect at least 1 resistant colony}) = p \qquad (1)$$

It is clear therefore that, as Equation (1) has no dependence upon b (the number of colonies tested) that there is no additional benefit to the probability of detection in testing more than one bacterial colony from an individual sample.

When n faecal samples are examined, the probability of detection within the pen becomes

$$Pr(\text{detect resistance} \mid \text{test } n \text{ individual faecal samples})$$

$$= Pr\ (\text{at least 1 faecal sample tests resistant})$$

$$= 1\text{- } Pr \text{ (none test resistant)}$$

$$= 1\text{- } Pr \text{ (1st sample tests sensitive} \cap \ldots \cap n^{th} \text{ sample tests sensitive)}$$

$$= 1\text{-}(1\text{-}p)^n \quad (2)$$

Extending to sampling n faecal samples from each of d pens, where d is the total number of pens sampled, the within-farm probability of detecting at least one resistant sample is given as: Pr(detect resistance| test n individual faecal samples from each of d pens)

$$= 1\text{-}(1\text{-}p)^{nd} \quad (3)$$

Composite Sample Model

Again assume faecal samples from n pigs in a single pen are collected, as illustrated in Fig. 1. A faecal sample is collected from each of the n pigs and then pooled, in equal measures, to form a composite sample representing the faecal constituent of the pigs in the pen. The composite sample is then homogenised and an inoculate formed, implying that the organisms are uniformly distributed. From this sample, m sub-samples are formed and tested for *Salmonella*. From each *Salmonella* sample, b bacterial colonies then are selected randomly and tested to classify a colony as sensitive or resistant to the antimicrobial considered. The type of susceptibility test used may vary according to the type of surveillance being carried out, but here it is assumed that the disc diffusion method is used. As the composite may contain both sensitive and resistant organisms, the probability of detecting resistance is dependent both on the number of sub-samples which are formed from the composite, m, and the number of bacterial colonies tested per sub-sample, b. Some surveillance schemes, in both the UK and other EU countries, test for resistance only one sub-sample of faeces per composite sample and one *Salmonella* colony per sub-sample. However the models developed here are generic in nature and so can be adapted easily to account for different values of m and b.

Assuming that the test is perfect and always produces the correct antimicrobial susceptibility classification, and that b *Salmonella* colonies are sampled from a single sub-sample for the detection of resistance, at least one colony must show resistance to the antimicrobial of interest to confirm resistance. Assuming that the tests conducted on successive colonies are independent, the probability that a chosen colony is classified as resistant is simply the prevalence of resistant bacteria within the composite sample, i.e. the proportion of all bacteria that are classified as resistant (denoted p_c). The argument when b colonies are tested is as follows

$$\text{Pr(detect resistance in sample | test } b \text{ colonies)}$$

$$= \text{Pr (at least 1 colony tests resistant)}$$

$$= 1\text{- Pr (none test resistant)}$$

$$= 1\text{- Pr (1st colony tests sensitive} \cap \ldots \cap b^{\text{th}} \text{ colony tests sensitive)}$$

which, assuming that each test is independent, simplifies to

$$= 1\text{- } Pr \text{ (1st colony tests sensitive)} \ldots Pr \text{ (}b^{\text{th}} \text{ colony tests sensitive)}$$

$$= 1 - \prod_{i=1}^{b} (1 - p_c)$$

$$= 1 - (1 - p_c)^b \quad (4)$$

Following the same approach the probability of detecting resistance in the pen is dependent on finding at least one resistant *Salmonella* sub-sample from the composite sample. Given that *m* sub-samples are tested per composite sample the probability of detection is

$$\Pr(\text{detect resistance} \mid \text{randomly test } m \text{ sub-samples})$$

$$= 1 - \Pr(\text{all } m \text{ sub-samples test sensitive})$$

$$= 1 - \prod_{i=1}^{m} (1 - p_c)^b$$

$$= 1 - (1 - p_c)^{bm} \quad (5)$$

Considering detection at the farm level, the probability of detecting resistance when *c* composite samples are formed is given as

$$\Pr(\text{detect resistance} \mid \text{randomly test } c \text{ composite faecal samples}) = 1 - (1 - p_c)^{bmc} \quad (6)$$

In summary, the probability of detecting antimicrobial resistant *Salmonella* for both composite and individual samples at the faecal, pen and farm levels is dependent upon the number of tests conducted and summarised in Table 1.

Table 1. Probability models for the detection of resistance

Sampling level	Individual	Composite
Faecal	P	$1-(1-p_c)^b$
Pen	$1-(1-p)^n$	$1-(1-p_c)^{bm}$
Farm	$1-(1-p)^{nd}$	$1-(1-p_c)^{bmc}$

Accounting for misclassification errors

The probability models in Table 1 assume that the diagnostic power of the test is perfect. Here, it was assumed that the antimicrobial susceptibility tests are conducted by a method known as zone disc diffusion. Misclassification errors can result from both experimental error and incorrect placement of the breakpoint concentration used to classify resistance relative to the minimum inhibitory concentration (MIC) distributions of truly sensitive and resistant strains. Figure 2 indicates how the rates of misclassification error can be calculated experimentally. Using the standardised disc diffusion technique to obtain the zone diameter for various antimicrobial concentrations and plotting these data against MIC data derived using a gold standard test gives a scatter plot as indicated in Fig. 2. Using these data the breakpoint can be adjusted so that the rates of false sensitive (*FS*) and false resistant (*FR*) reporting are as small as

possible. The British Society for Antimicrobial Chemotherapy (BSAC) working party considers it acceptable to have the *FR* rate <5% and the *FS* rate <1% (MacGowan & Wise, 2001). This cautious approach puts tighter constraints on the reporting of false sensitives as these are of greater clinical importance due to possible treatment failure.

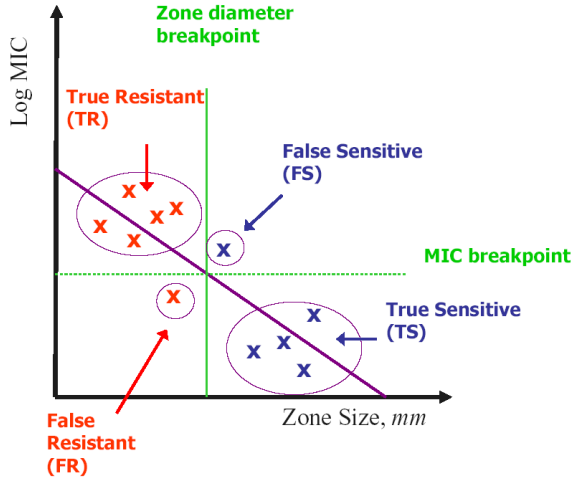


Fig. 2. Illustration of the possible misclassification of resistance status when using zone disc diffusion

Misclassification errors occur due to reduced test sensitivity (*Se*) or specificity (*Sp*). The sensitivity of a test for resistance is the proportion of truly resistant isolates detected by the test. The specificity of the test is the proportion of truly sensitive isolates detected. In terms of Fig. 2 these terms can be calculated as $Se = \frac{TR}{TR + FS}$ and $Sp = \frac{TS}{TS + FR}$. Using the BSAC Working Party's acceptable rates for *FR* and *FS* reporting, the corresponding recommended values of test sensitivity and specificity are therefore $Se > 0.99$ and $Sp > 0.95$. The probability models, summarised in Table 1, can be adapted to account for imperfect test sensitivity (*Se*) and specificity (*Sp*). Such adaptation leads to farm level probability of detection (*PD*) models of

$$PD=1-[p(1-Se)+(1-p)Sp]^{nd} \quad (7)$$

for individual level samples and

$$PD=1-[p_c(1-Se)+(1-p_c)Sp]^{bmc} \quad (8)$$

for a composite sample.

Sample size and minimum detectable difference

Manipulating Equations (7) and (8), the total number of tests required to detect antimicrobial resistance with a specified probability *PD* can be calculated as

$$nd = \frac{\log(1-PD)}{\log(p(1-Se)+(1-p)Sp} \tag{9}$$

for the individual model, and similarly a total number of tests

$$bmc = \frac{\log(1-PD)}{\log(p_c(1-Se)+(1-p_c)Sp)} \tag{10}$$

for the composite model. Further rearrangement also allows the minimum detectable prevalence to be calculated, as shown in Equations (11) and (12) given that the required probability of detection and total number of tests are defined, as

$$p = \frac{1}{(1-Se-Sp)}\left[\exp\left(\frac{\log(1-PD)}{nd}\right)-Sp\right] \tag{11}$$

and

$$p_c = \frac{1}{(1-Se-Sp)}\left[\exp\left(\frac{\log(1-PD)}{bmc}\right)-Sp\right] \tag{12}$$

for the individual model and composite model, respectively.

Quantifying the variability associated with *Se* and *Sp*

Although the sensitivity and specificity of the test procedure may have been established according to the BSAC guidelines, it is possible that the values of *Se* and *Sp* vary, for example, between laboratories, days and technicians. This variability may be due to differences in test procedure associated with these factors, although sampling protocols, testing protocols and staff training should keep this variation to a minimum. This variation is characterised by describing *Se* and *Sp* as normal random variables, that is $N(\mu_{Se},\sigma^2_{Se})$ and $N(\mu_{Sp},\sigma^2_{Sp})$, respectively. Expressing the variability in this manner allows 95% intervals to be calculated for the probability of detection. In order to do so, estimates of μ_{Se}, μ_{Sp}, σ_{Se} and σ_{Sp} are required. It is known that the required values of *Se* and *Sp* according to BSAC guidelines are *Se*=0.99 and *Sp*=0.95. Assuming, therefore, that this is attainable on average, it can be assumed that μ_{Se}=0.99 and μ_{Sp}=0.95. The likely value of the standard deviation, σ, is however, unknown. It is therefore assumed that the data are normally distributed and so 99.7% of the data are contained within three standard deviations (3σ) of the mean value. Assuming that 99.7% of the variability can be represented as δ% of μ_{Se} or μ_{Sp} an expression for the value of σ are $\sigma_{Se} = \frac{\delta\mu_{Se}}{3}$ and $\sigma_{Sp} = \frac{\delta\mu_{Sp}}{3}$ where δ is chosen systematically. The probability of detection is then simulated for the chosen value of δ. Repeat simulations allow for calculation of the mean, median, upper (0.975) and lower (0.025) percentiles of the probability of detection. One thousand iterations ensures that the 95% variability interval surrounding the probability of detection is accurately calculated.

Model parameterisation

Results are presented for all possible values of within-pen prevalence, p, when the total number of tests (*nd* for individual samples) is 1, 2, 4 and 8. Various values of prevalence are considered in the model, but for certain analyses it is assumed that p=0.01, which focuses on the probability of detection when resistance is emerging. The number of tests required to detect resistance at this prevalence is established when *Se* and *Sp*=1 and when *Se*=0.99 and *Sp*=0.95 (BSAC guidelines) for three different probabilities of detection *PD*=0.99, 0.95 and 0.90.

Again, for low prevalence p=0.01, the variability of *Se* and *Sp* is examined using values of δ=0.1 and δ=0.2. Using δ=0.1 implies that 99.7% of the simulated values of *Se* lie in the range 0.957-1 and the values of *Sp* in the range 0.918-0.98, and using δ=0.2 results in a range of 0.924-1 for *Se* and 0.886-1 for *Sp*.

Finally the minimum detectable within farm prevalence is calculated, when *PD*=0.99, 0.95, 0.90 and 0.50 when the total number of tests equals 1, 2, 8, 10 and 100.

RESULTS

The results presented are given for the individual sampling model where p is defined as the prevalence of resistant infection within the pen. However, the results are identical for the probability of detection at the farm level for the composite sampling model if the values used for p are applied to p_c (proportion of resistant organisms within the composite sample) and *nd* = *bmc*. However it is important to note that p and p_c are not equal. This is because p_c is dependent on the proportion of animals in a pen that are infected with antimicrobial sensitive *Salmonella*. A low value of p_c would occur if few animals infected with resistant *Salmonella*, but many with antimicrobial sensitive *Salmonella* contributed to the composite sample.

Probability of detection at the farm level

Assuming a perfect test, Fig. 3 illustrates that, when resistance is emerging and hence the within pen prevalence is at a low level, increasing the number of tests (*nd*) leads to increases in the probability of detection at the farm level. As the prevalence of resistance increases, the benefit of additional tests diminishes, for example when p=0.7 increasing the number of tests from 4 to 8 does not increase the probability of detection.

For low prevalence (p=0.01), the number of tests required to detect resistance with 95% certainty (*PD* = 0.95), is 299 as shown in Table 2. Extending these results to a test with *Se*=0.99 and *Sp*=0.95, the BSAC working group recommended level, only 49 samples are required to be 95% confident of detection when p=0.01. This differential occurs because some antimicrobial sensitive isolates are misclassified as resistant because of the reduced test specificity. Increasing the required probability of detection to 99% requires around a 50% increase in the number of tests needed.

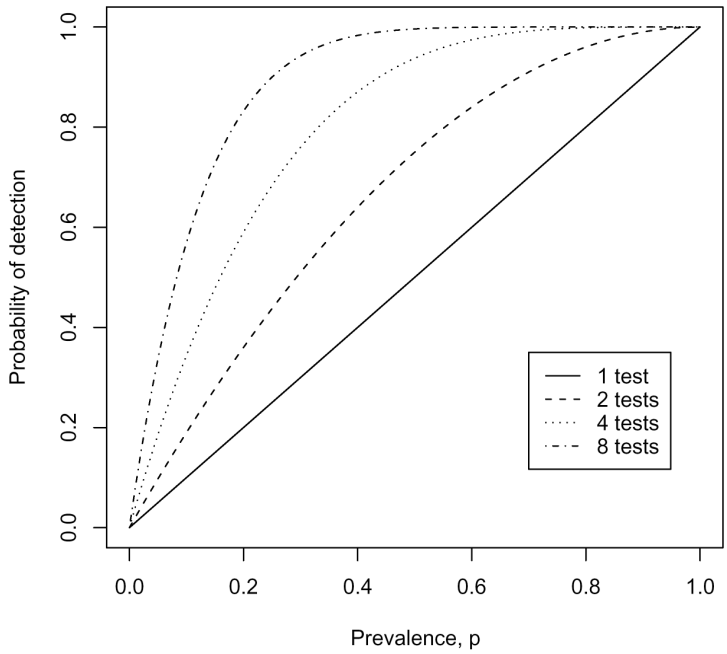


Fig. 3 The probability of detecting resistant *Salmonella* for all possible values of pen level prevalence for a varying number of total tests (total tests = *nd* for individual samples).

Table 2. Number of tests (*nd*) required to detect resistance at the farm level where the pen level prevalence, p=0.01.

Probability of Detection (*PD*)	Number of tests required to detect resistance	
	Se, *Sp*=1	*Se*=0.99, *Sp*=0.95
0.99	459	76
0.95	299	49
0.90	230	38

Table 3 summarises the variability in the number of tests required to detect resistance when the sensitivity and specificity of the test are variable. It is clear that 1000 simulations accurately simulates the variation present, as the median estimate is equivalent to the deterministic estimate when *Se*=0.99, *Sp*=0.95. Defining the variability in *Se* and *Sp* using δ=0.1, implies the possible sample size may be as small as 23 or as large as 304 to detect resistance with 95% certainty, and the required sample size as large as 467 to be 99% certain of detection. Increasing the value of δ, as illustrated with δ=0.2 in Table 3, leads to an increase in the range of the sample size required.

Table 3. The number of tests (*nd*) required to detect resistance at the farm level when the pen level prevalence, p=0.01 and $Se \sim N(\mu_{Se}, \sigma^2_{Se})$, $Sp \sim N(\mu_{Sp}, \sigma^2_{Se})$ where μ_{Se}=0.99, μ_{Sp}=0.95, $\sigma_{Se} = \frac{\delta Se}{3}$ and $\sigma_{Sp} = \frac{\delta Sp}{3}$.

Probability of detection (*PD*)	Number of tests required to detect resistance median (lower percentile (2.5%), upper percentile (97.5%))	
	δ=0.1	δ=0.2
0.99	76 (35, 467)	76 (24, 504)
0.95	49 (23, 304)	49 (16, 327)
0.90	38 (19, 235)	38 (12, 253)

Minimum detectable prevalence

The results for the minimal detectable prevalence when using a diagnostically perfect test conducted for 1, 2, 8, 10 and 100 samples are outlined in Table 4. These results illustrate that conducting a single test has the ability to detect a within-pen prevalence only equal to the probability of detection. Increasing the number of tests to, for example, 8 allows detection when p=0.44. In order to have a high probability of detecting a lower prevalence the number of tests must be substantially increased. A probability of detection (*PD)* of 0.99 requires 100 tests to detect a prevalence (*p)* of 0.05. A lower prevalence is detectable if the required probability of detection is lowered, which however implies that there would be more occasions when the resistant infection is not found.

Table 4. The minimum detectable prevalence when the number of tests and the required probability of detection on farm is varied, *Se*, *Sp*=1

		Probability of Detection *(PD)*			
		0.99	0.95	0.90	0.50
Number of tests	1	0.99	0.95	0.90	0.50
	2	0.90	0.78	0.68	0.29
	8	0.44	0.31	0.25	0.08
	10	0.37	0.26	0.21	0.07
	100	0.05	0.03	0.02	0.01

DISCUSSION

Information relating to the levels of antimicrobial resistant infections in livestock has become more important in recent years due to the potential exposure of the public to an additional pool of resistant organisms via the food chain. The probabilistic models presented in this paper present a means of providing insights on the likelihood that antimicrobial resistant *Salmonella* is being detected on a farm under typical survey systems. Use of such models may be particularly beneficial for the design of active surveillance programmes and research projects.

Assuming a perfect test, modelling the probability of detection showed, as expected, that increasing the number of tests (*nd* for individual samples, *bmc* for composite sample), whether increasing the number of pens tested (*d*) or the number of composites formed (*c*), increases the probability of detection at the farm level. This is particularly true when resistance is present at a low prevalence. This result shows that testing a single isolate/colony from a herd (*b,m,c*=1 for a composite and *n,d*=1 for an individual sample) requires a high prevalence of resistance within the pen or within the composite sample to detect resistance (i.e. p=0.95 or p_c=0.95 when *PD*=0.95). This approach is likely to be sufficient if the farm and pen are clonally infected i.e. all of the *Salmonella* organisms present within the farm are identical, with the same resistance profile. From passive surveillance work for *Salmonella* in Great Britain (VLA, 2007), it is known that there is usually one *Salmonella* serovar present and that sometimes the antimicrobial resistance pattern is intrinsic to the serovar e.g. *Salmonella* typhimurium DT104. However, if this is not the case, then this strategy may not be sufficient to detect emerging resistance on the farm.

The tests used to detect antimicrobial resistance are unlikely to be perfect. Using the BSAC recommended values of *Se*=0.99 and *Sp*=0.95, the model predicted that fewer tests were required to detect resistance than for a perfect test (*Se* and *Sp*=1). This counter-intuitive result arises because some sensitive samples are misclassified as resistant.

The idea of the imperfect test was explored further by modelling the variation in *Se* and *Sp* of antimicrobial susceptibility testing associated with differences between laboratories, days and technicians. Results showed that the detection of resistant organisms is highly dependent on the values of *Se* and *Sp*. For a fixed value of within-pen prevalence, p, or the proportion of resistant organisms within the composite p_c, as *Se* and *Sp* become more variable the percentile bands for the number of tests required also become wider. The effect of this widening is, however, dependent on the within-pen prevalence or the proportion of resistant organisms within the composite. For low p or p_c, *Sp* has a strong influence and conversely for high p or p_c, *Se* has a strong influence. Increasing the levels of *Se* and *Sp* for low p or p_c increases the probability of detection but also leads to a decrease in the probability of a correct diagnosis. This indicates that, if the values of *Se* and *Sp* achieved in a laboratory are very variable, the results can be misleading. Generally, efforts are made to limit variation between laboratories by adopting standardised testing methods, by limiting the number of laboratories that carry out the tests and by carrying out ring-trials with other laboratories.

As mentioned previously, the number of tests conducted has a large effect on the probability of detection. Conducting an analysis on the required sample size for a defined level of detection on farm showed that, for a perfect test, a large sample size (299 tests) is required to detect low level resistance (p=0.01 or p_c = 0.01) with 95% certainty. This sample size was found to diminish exponentially with prevalence if the test were assumed to meet BSAC requirements

only (*Se*=0.99, *Sp*=0.95). This counterintuitive result again shows the effect of misclassification leading to detection when the prevalence is low. Such misclassification errors are unpreventable. It is, however, important to understand why they occur and the effect they have before any results are interpreted. Conducting more tests increases the ability to detect lower prevalence, however the number required to detect resistance at a low prevalence accurately is high and is not likely to be practicable. This implies that the probability of detection of emerging (low prevalence) resistance on a farm is low if only one isolate per sample is tested. For low level resistance, it would be better to use selective plating (where the antimicrobial is in the plate at the specified breakpoint so that any growth confirms the presence of a resistant *Salmonella*).

In Great Britain, the (passive) system of testing clinically diseased animals is complemented by regular and sometimes ongoing structured (active) surveillance of healthy animals. Provision of evidence from both populations is very helpful. Passive surveillance for *Salmonella*, although not designed to detect new and emerging resistance may do so as this may be more likely to appear in diseased animals, perhaps because the resistance confers pathogenicity or because antimicrobial treatment selects for or induces resistance. However, if selective plating is not used, a high prevalence of resistance may be required for detection to occur.

In conclusion, the probability of detecting resistant *Salmonella* has been shown to be dependent upon the level of resistance within pen/farm (individual animal sampling) or within sample (composite sampling) and the variable diagnostic power of the test used. The likelihood of detecting low level resistance on individual farms, such as may be present for emerging resistance, was found to be low and therefore selective plating is recommended in such cases. The theoretical models developed and presented here provide an insight into the sensitivity of the sampling and testing methods (e.g. what sampling method should be used and how many tests should be conducted) and could be used to inform any future on-farm surveillance programmes or research projects.

REFERENCES

Arnold M., Cook A. and Davies R. (2005) A modelling approach to estimate the sensitivity of pooled faecal samples for isolation of *Salmonella* in pigs. Journal of the Royal Society, Interface. 2, 365-372.

Davison H.C., Low C.J. and Woolhouse M.E.J. (2000) What is antibiotic resistance and how can we measure it? Trends in Microbiology. 8, 554-559.

MacGowan A.P. and Wise R. (2001) Establishing MIC breakpoints and the interpretation of *in vitro* susceptibility tests. Journal of Antimicrobial Chemotherapy. 48, (Supplement 1) 17-28.

Swann M. M. (1969) Report, Joint Committee on the use of Antibiotics in Animal Husbandry and Veterinary Medicine, HMSO, London.

VLA (2007) *Salmonella* in Livestock Production in GB: 2006 Report. Veterinary Laboratories Agency.

A STRUCTURED JUDGEMENT STUDY ON *SALMONELLA* SPP. IN PORK: ANALYSIS OF DIFFERENT WEIGHTING SCHEMES

I. BOONE*, Y. VAN DER STEDE, K. BOLLAERTS, W. MESSENS, K. GRIJSPEERDT, G. DAUBE AND K. MINTIENS

SUMMARY

A structured expert judgement study was set up in order to complete missing input parameters for a Quantitative Microbial Risk Assessment (QMRA) model to estimate the risk of human *Salmonella* infections associated with the consumption of minced pork. A performance-based weighted linear pooling model, known as Cooke's Classical Model was used. The experts' judgements for all the estimated variables, expressed by subjective probability density functions, were weighted and then combined into one "decision maker" (DM) distribution. Weights were calculated based on the experts' performance on calibration questions. The aim of the study was to compare four different weighting schemes involving item weight, global weight, equal weight and user weight DM. The item weight DM obtained the highest performance and outperformed the other DMs. The scores for self-rating of expertise were generally not indicative of the experts' performance on the calibration variables.

The applied classical model provided rational basis to select a weighting scheme to compute each of the variables of interest as input in the QMRA model. Attention is necessary to find adequate and relevant calibration variables, since this is important for the validation of the results of the weighting scheme.

INTRODUCTION

Very often decision makers cannot wait until the risk analysts obtain all the necessary empirical data to complete a model. Expert judgement is frequently used to provide input for Quantitative Microbial Risk Assessment (QMRA), when empirical data are missing, difficult to obtain or in case of poor quality (Alban & Stark, 2005; Nauta et al., 2001; Stark et al., 2002; Van der Fels-Klerx et al., 2005). As with empirical data, the input obtained through expert opinion may be biased and imprecise and can also have an impact on the outcome of the risk assessment (Walker et al., 2003). Walker and co-workers defined the quality of an expert judgement as how accurate the judgement estimates the true value and how well it is related to what the expert knows about the subject. Since the quality of expert judgement may ultimately determine the validity of decisions based on a QMRA, it is important that the expert judgement's are elicited and treated through a formal approach using transparent and objective methodological rules. Different methods involving both mathematical and behavioural

* Ides Boone, Veterinary and Agrochemical Research Centre, Coordination Centre for Veterinary Diagnostics, Groeselenberg 99, 1180 Brussels, Belgium. Email: ides.boone@var.fgov.be

approaches are used to elicit expert judgements and to combine experts' subjective probability distributions (Clemen & Winkler, 1999; Cooke, 1991; Scholz & Hansmann, 2007).

A structured expert judgement study (Cooke, 1991) was conducted in order to fill in missing input parameters (variables of interest) for a QMRA-model. The purpose of this model was to evaluate the risk of human illness associated with the consumption of minced pork meat due to *Salmonella* spp. in Belgium. The missing input parameters were related to (i) the bacterial contamination load and prevalence of pig carcasses during some key processing steps in the slaughterhouse, (ii) the agreement between bacteriological and serological test results related to *Salmonella* in pigs (iii) the effect of transport, holding and lairage on the excretion of *Salmonella* spp. in pigs. The experts' opinions on the variables of interest expressed as subjective probability distributions were combined into a single combined distribution, using Cooke's classical model, which is a performance-based weighted averaging model (Cooke, 1991).

This model was chosen in this study because it aims at enhancing a rational consensus on the combined experts' distributions of the missing input parameters, by complying with four necessary conditions (Cooke, 1991): 1) the whole process should be open for peer review, 2) expert's assessment must allow for empirical quality controls, 3) the elicitation procedure should encourage the experts to state their true opinion and not bias the results, and 4) the experts should not be pre-judged before processing the results.

In the classical model, the experts' weights are derived from two quantitative performance measures namely calibration and information. Both satisfy a proper scoring rule and can be obtained by using seed (calibration) variables. Seed variables are variables whose true values are known to the analyst but unknown to the experts or which will become known *post-hoc*. The performance of the experts to the seed variables is taken as indicative to their performance for the specific variables of interest. Seed variables are used to serve three objectives: i) measure expert performance, ii) enable performance-based weighted combination of experts' distributions and iii) evaluate the combination of the resulting expert opinions (or decision maker) (Cooke & Goossens, 2008). Using these seed variables we are able to classify experts according to their calibration and information scores. Experts who made better assessment regarding the seed variables obtain higher calibration scores (close to one) resulting in larger influence in the decision maker's distribution (Vargas Galindo, 2007).

A calibration score can be interpreted as the likelihood of the hypothesis that an expert's probabilistic statements are correct and that a set of experimental results (realisations to seed variables) correspond with the experts' assessment. Information scores measure the degree to which a distribution is concentrated, in other words how small the experts' uncertainty bounds are (Cooke & Goossens, 2006). Information is relative to a background measure and is measured for both the seed variables and variables of interest. For each variable, the background measure (uniform or log-uniform) is taken as the smallest interval containing all the assessed quantiles and the realisation, if available, of all experts, augmented and decreased with 10%. Larger information scores are obtained when the experts elicit quantiles that are located closely together.

The purpose of this paper was to evaluate the quality of combined expert's distributions by comparing performance-based decision makers with the equal weight decision maker. The equal-weight decision maker assigns equal weight to each expert and does not take into account the experts' performance on the seed variables. In addition, a user weight decision maker was

constructed by using weights related to the experts' self-assessment of expertise which was filled in during the questionnaire. The correlation between the self-rating of the experts and performance-based scores was analysed to assess for possible over/under confidence of the experts. By comparing the different weighting schemes, we hypothesized that the resulting performance-based decision maker outperformed equal weighting, and the decision-makers' distribution outperformed the best expert.

MATERIALS AND METHODS

Elicitation protocol

The elicitation protocol used was based on methods described by Cooke and Goossens (2000) and van der Fels-Klerx et al. (2002; 2005). It consists in the preparation for elicitation (definition of case structure, identification of variables of interest and seed variables, identification of experts, design of quantitative elicitation session and dry-run session), the elicitation itself and the post-elicitation (combination of the experts' assessments, robustness analysis and documentation of the results).

The elicitation documents consisted of two separate questionnaires which were merged. A first part was related to the background of the experts, a second part to specific questions on gaps in the QMRA-model pertaining to the epidemiology of *Salmonella* spp. in the pork production chain. In total six and 18 questions were related to respectively variables of interest (Table 1) and seed questions (Table 2). The seed variables included questions related to *Salmonella* prevalence at various stages in the slaughterhouse, to production parameters in Belgian slaughterhouses and cutting plants, to the number of salmonellosis cases in the EU and questions related to the bacteriological diagnostic test (ISO 6579: 2002) used in Belgium.

The true values or realisations for the seed variables were obtained from published and unpublished data. An expert was considered to be a professional involved in pork meat supply chain with an advanced knowledge of the epidemiology and/or microbiology of *Salmonella* in pigs and/or pork. A list of delegates of the 7th International Symposium on the epidemiology and control of foodborne pathogens in pork (Safepork, 9-11 May, 2007, Verona, Italy) was obtained from the conference organisers. The delegates that submitted an abstract on a *Salmonella* topic were contacted (61 persons). An additional nine persons were also contacted when a search on PubMed revealed they had several relevant publications.

The questionnaire was pre-tested by two Belgian experts, not involved in the expert elicitation workshop (dry run session).

Table 1. Variables of interest

ID	Name
Question 1	*Salmonella* concentration at following abattoir steps :
V[a]1	Unloading pigs from truck at the abattoir
V2	Lairage
V3	Stunning and killing
V4	Scalding
V5	De-hairing
V6	Singeing
V7	Polishing
V8	Evisceration
V9	Splitting
V10	Meat Inspection
V11	Chilling
Question 2	*Salmonella* bacteriological prevalence (%) in subsequent steps at slaughterhouse (starting with a 7% prevalence when pigs are leaving the farm)
V12	Unloading at the abattoir
V13	When euthanized
V14	After singeing
V15	After polishing
V16	After evisceration
V17	After meat inspection
V18	Chilled carcass
Question 3 (V19)	% Bacteriological positive pigs given these pigs are serologically positive (positive test agreement bacteriology and serology given serologically positive)
Question 4 (V20)	Percentage of pigs excreting *Salmonella* spp. after transport and lairage in the slaughterhouse (starting with a *Salmonella* excretion percentage of 5% when pigs are leaving the farm)
Question 5 (V21)	Increase of *Salmonella* spp. prevalence among pig carcasses due to improper cleaning of the conveyor belt and work surface at the cutting plant
Question 6 (V22)	Relative contribution of *Salmonella* Typhimurium human salmonellosis cases due to the consumption of minced pork (versus non-minced pork)

[a]: ID-number for variable of interest

Table 2. Seed variables

ID	Name	Realisation
S[a]1	Within-herd apparent sero-prevalence (in %) for *Salmonella* in Belgian pig herds in 2006.	36.8
S2	Percentage of serological samples found positive for *Salmonella* spp in Belgian pigs sampled in 2005.	39.25
S3[b]	Incubation time for the *Salmonella* detection test ISO 6579: 2002 during pre-enrichment in Buffered Peptone Water	18 (ISO 6579: 2002)
S4[b]	Ideal pH of water in pigs' drinking water at the farm in order to reduce *Salmonella* spp. after adding acids to the water	3.9
S5	Average duration for fasting pigs before slaughtering (in hours) in Belgium	17
S6	Average number of slaughtered pigs per hour in 10 biggest slaughterhouses in Belgium	398.5
S7	Average duration of pigs kept at the lairage in the slaughterhouse in Belgium (minutes)	126
S8	Duration of the singeing process (in seconds) in the slaughterhouse	10.6
S9[b]	Average pig carcass weight in Belgium (kg)	82.5
S10	*Salmonella* bacteriological prevalence through swabbing of 600 cm² of a pig's carcass in five large slaughterhouses a) after polishing	11.1
S11	b) after splitting	13.7
S12	c) of chilled carcasses	2.2
S13[b]	Minimum growth temperature for *S.* Typhimurium in pig meat	9
S14	Average temperature in the working hall of the 11 largest Belgian cutting plants (°C)	9.6
S15[b]	Duration (in minutes) for manipulation of pig carcasses in the working hall in Belgian cutting plants (half carcasses are cut into shoulder, back, belly, ham, and into smaller pieces)	38
S16	Average number of pig meat servings per person per year in Belgium	11
S17[b]	Number of salmonellosis cases in EU in 2005	176,395 (EFSA, 2006)
S18[b]	Number of salmonellosis cases in EU in 2006	NA[c]

[a]ID-number for seed variable; [b] seed variables discarded from the analysis; The realisations of the seed variables retained in the study originated from unpublished Belgian data, mostly collected during the METZOON research project. [c]NA: realisation not available yet

A workshop was organised at the end of the first day of the conference. The Belgian QMRA model was orally presented in a plenary session earlier during the conference. A short introduction was given to the participants of the workshop on how the questionnaire should be filled in. In order to represent the uncertain quantities of the variables by a subjective probability density function (PDF), the experts were asked to give a three-point estimate. Hereto, a most likely value, a minimum and a maximum value were asked for each seed variable and variable of interest (van der Gaag & Huirne, 2002; Vose, 2000). For each question, the expert's self-rating of his experience was asked.

All experts had to complete the questionnaire individually. The second questionnaire (13 questions) was send later by e-mail to a selection of experts having completed an adequate

number of questions in the first questionnaire. The answers to these questions were individually filled in by the experts and returned by e-mail.

Before the onset of the analysis, the answers were screened for inconsistencies and verified for misunderstanding of questions. In the aggregation of the experts' subjective PDFs, weights were computed using the Classical Model which has been implemented in the Excalibur software (Pro-version v 1.0, developed by TU Delft and provided by R.M. Cooke; EXCALIBUR light version available at http://dutiosc.twi.tudelft.nl/~risk/) in order to obtain one combined PDF. The Excalibur software allows parametric and quantile input from continuous uncertain quantities elicited by experts and combines these according to the methods of R.M. Cooke (Cooke, 1991). Estimates for both the seed and variables of interest (a minimum, a most likely and a maximum value) were modelled as a triangular distribution. These parameters were introduced in EXCALIBUR, and the according 5%, 50% and 95% quantiles for these triangular distributions were obtained.

The classical model combines the calibration and information score into a single overall combined score, in which the calibration score dominates over the information score in the calculation of the decision maker. Using the classical model, different combination schemes to obtain the decision makers' distribution can be compared. For a detailed description of the scoring rules, we refer to (Cooke, 1991; Cooke & Goossens, 2000; Lopez De La Cruz, 2004).

The equal weight decision maker results from the weighted sum of the experts' individual distributions, with weights taken equal for all the experts. When there are N experts, the weights for each PDF equal 1/N. If N experts have assessed a given set of variables, the equal weight decision maker's distribution is given by Equation 1:

$$f_{eqdm,i} = \left(\frac{1}{N}\right)\sum_{J=1}^{N} f_{j,i} \ , \tag{1}$$

where $f_{j,i}$ is the density associated with experts j's assessment for variable i.

There are two performance-based weighting techniques available in EXCALIBUR, to combine the experts PDF's based on the experts' performance on the seed variables: the "global weight decision maker" and the "item-weight decision maker". Both weighting schemes satisfy a proper scoring rule (Cooke, 1991). In the global weight decision maker the weights are defined by the normalized product of the calibration and the overall information scores on the seed variables, i.e. one set of weights for all questions (Eq. 2). For variable i, the global weight decision maker's density is:

$$f_{gwdm,\,i} = \sum_{j=1}^{N} w_j f_{j,i}; \sum_{j=1}^{N} w_j = 1 \quad , \tag{2}$$

where w_j is the normalised weight of the j^{th} expert and $f_{j,i}$ as in Equation 1.

In the item weight decision maker, the weights are determined per expert and per variable, using the experts' information score for each variable, instead of using an overall measure of

information as in the global weight decision maker (Eq. 3). The item weight decision maker's density for variable i is given by:

$$f_{iwdm,\,i} = \sum_{j=1}^{N} w_{j,\,i} f_{j,\,i}; \sum_{j=1}^{N} w_{j,\,i} = 1 \qquad (3)$$

The performance-based decision makers are optimised as follows: when a calibration score for an expert falls below a certain cut-off level ($\alpha > 0$), this expert should be attributed a zero-weight. The weights of the remaining experts will normalize to 1. With each value of α, a decision maker (DMα) is defined, computed as a linear combination of the experts whose calibration score exceeds the cut-off α. The weight of DMα equals its weight as if it would have been added as a virtual expert. The value of α is obtained by optimisation: the cut-off value for α for which the DMα receives the highest weight is chosen as cut-off value for determining the unweighted (weight "0") experts.

Next to equal weighting and performance-based weighting, the user-weight decision maker attributes weights to the experts which are imputed by the analyst. The user-weights given were computed based on the experts self-rating scores, by asking to asses his/her personal level of expertise per parameter (on a five-point scale from no knowledge to high expertise). These self-assessment scores were used to evaluate over/under confidence of experts based on their calibration scores using the Classical Model and to compute the user-weight decision maker.

A robustness analysis was performed to check how the results would change if different experts or seed variables were used.

RESULTS

Participation

Out of the 70 participants at the symposium, twenty-seven experts, originating from nine European countries (Belgium, Czech Republic, Denmark, Germany, Ireland, Norway, Spain, the Netherlands, the UK), the USA and Canada, completed the first questionnaire. Fourteen of these attended the workshop. Moreover, ten experts who could not attend the workshop submitted the questionnaire later on during the symposium. Finally three experts sent their completed questionnaire by mail or by e-mail shortly after the symposium.

Based on first questionnaire's completion rate, a second questionnaire containing the majority of the seed variables was returned by 11 experts (on 21 experts) originating from seven European countries (Denmark, Germany, Ireland, Norway, Spain, the Netherlands, the UK), the USA and Canada. Due to the limited number of experts per country (on average one expert per country) a country effect could not be studied. The experts covered all the areas in the pork production chain, but experts active in the animal production domain and/or in combination with animal transport, holding and slaughter (n=7) and in the consumer and public health part (n=5) were predominant. Two experts indicated they were active in the retail and distribution area. Seven experts worked in a research institute, and indicated their field of interest was primarily epidemiology (n=10) and antimicrobial resistance (n=9). Seven experts peer reviewed publications related to *Salmonella* in the pork production chain.

Performance of experts

In total, eleven effective seed variables were used to weigh the experts' opinions since that was the minimum number of seed variables assessed by 11 experts. From the initial 18 seed variables, seven seed variables were dropped (Table 2): Two variables were ambiguously defined (S4 and S13). Pre-knowledge of the realisation by one of the experts was obvious for one seed variable (S17), while from another seed variable the true value was not available during the timeframe of the study (S18). The experts' elicitation to three seed variables (S3, S9 and S15) was missing for one of the 11 experts.

Table 3 presents the numbers of target and seed variables, the performance measures of the performance-based decision makers, the equal weight DM, user weight DM and the best expert's performance. The item weight DM obtained the highest calibration and information scores as compared to the global, equal and user weight DMs. It appeared that the item-weight DM would provide the best estimates for the variables of interest. In the optimization procedure of the performance-based DM (global weight DM and item weight DM) weight was allocated to three experts (4, 9 and 11) while the others experts were unweighted. The user weight DM had the lowest calibration and information scores. The equal weight and user weight DM obtained significantly lower information scores than the performance-based DM. In general, the information score for the calibration variables were lower than for the variables of interest. This was most pronounced with the best expert (No. 11) where the information score was 0.615 and 1.235, respectively.

Table 3: The performance of the virtual expert, applying item weights, global weights, user weights and the best expert, and the number of target variable vs. effective seed variables.

Target/seed variables	Performance measure	Virtual expert				
		Item weights	Global weights	Equal weights	User weights	Best Expert (No.11)
21/11	Calibration	0.6150	0.3851	0.4920	0.3697	0.2151
	Information	0.5423	0.4449	0.2443	0.2406	0.6657
	Combined weight	0.3335	0.1713	0.1202	0.0889	0.1432

Robustness analysis

Robustness analysis was used to identify the importance of each expert with relation to the decision maker. One expert at the time was excluded and then the relative information and calibration scores of the remaining experts were computed. The calibration scores of the item weight DM were lower if expert 11, 4 and 9 were removed (DM calibration score were 0.04576, 0.2151 and 0.3851, respectively). This indicates that these experts contributed significantly to the results. Expert 11 was the expert who obtained the highest weight, removing him/her from the pool of experts resulted in a considerable loss in the calibration score of the item-weight DM. Robustness against the choice of experts was higher in the global weight decision maker than in the item weight DM, since only removal of expert No. 4 and 11 resulted in a small loss of the global-weight DM calibration score. A robustness analysis carried out on the choice of the seed variables, indicated that the there were no seed variables which had a strong influential effect on the calibration score of the item-weight DM or the global weight DM.

Aggregated distributions

Seed variables: The combined DM distributions for the 11 seed variables are shown in Table 4, as expressed by their 5%, 50% and 95% quantile The realisations of ten variables fell within the 90% confidence interval of the item-weight DM distribution, whereas the realisation of seed S16 fell outside the 90% confidence interval (Table 4). The confidence intervals of the item weight decision maker were narrower in 4/11 seed variables as compared to the global weight decision maker.

Table 4a. Uncertainty distributions expressed by their 5%, 50% and 95% quantiles for the seed variables 1, 2, 5, 6, 7) with their true values (realisations) obtained from the item weight, global weight, equal weight, user weight DM as well as the best expert.

ID	DM	5%	50%	95%	REALISATION
S1	1	19.52	43.96	68.92	36.8
	2	15.64	42.96	68.97	
	3	9.66	36.27	68.38	
	4	11.04	39.49	68.29	
	5	27.75	45.36	69.05	
S2	1	10.71	38.96	68.51	39.25
	2	11.46	39.99	68.60	
	3	8.91	27.27	78.22	
	4	9.09	29.48	73.37	
	5	27.75	45.36	69.05	
S5	1	5.27	15.83	28.45	17
	2	4.83	16.11	29.18	
	3	1.78	12.23	26.67	
	4	2.17	12.17	26.94	
	5	4.65	15.22	29.43	
S6	1	51.05	459.10	792.10	398.5
	2	56.92	535.40	799.20	
	3	43.35	302.60	725.70	
	4	48.58	316.40	742.30	
	5	394.90	600.00	805.10	
S7	1	57.11	146.00	1064.00	126
	2	63.91	418.90	1131.00	
	3	4.75	147.80	948.10	
	4	6.17	161.40	947.60	
	5	104.30	471.90	1134.00	

S: seed variable number; DM: decision maker, 1: item weight DM, 2: global weight DM, 3: equal weight DM, 4: user weight DM, 5: best expert (nr 11)

Table 4b. Uncertainty distributions expressed by their 5%, 50% and 95% quantiles for the seed variables 8, 10, 11, 12, 14) with their true values (realisations) obtained from the item weight, global weight, equal weight, user weight DM as well as the best expert.

ID	DM	5%	50%	95%	REALISATION
S8	1	6.96	19.58	49.93	10.6
	2	7.36	27.79	50.82	
	3	1.74	14.37	46.93	
	4	2.12	19.08	47.82	
	5	17.07	32.61	51.34	
S10	1	2.46	9.24	23.58	11.1
	2	2.46	9.38	23.60	
	3	0.00	8.85	34.80	
	4	0.00	8.64	34.85	
	5	2.45	10.25	23.75	
S11	1	2.51	10.22	23.75	13.7
	2	2.52	10.61	23.70	
	3	1.69	16.42	51.82	
	4	1.73	17.28	52.89	
	5	2.45	10.25	23.75	
S12	1	1.97	7.31	19.66	2.2
	2	1.96	7.41	19.66	
	3	0.66	8.12	41.84	
	4	0.96	8.46	39.50	
	5	1.94	8.42	19.76	
S14	1	5.56	10.43	17.82	9.6
	2	5.57	10.66	17.94	
	3	-6.88	13.60	23.44	
	4	-6.55	13.15	22.40	
	5	5.55	9.80	16.78	
S16	1	13.33	32.32	174.40	11
	2	13.52	37.26	338.60	
	3	10.67	84.68	312.30	
	4	10.53	80.75	312.90	
	5	13.29	31.28	50.92	

S: seed variable number; DM: decision maker, 1: item weight DM, 2: global weight DM, 3: equal weight DM, 4: user weight DM, 5: best expert (nr 11)

Variables of interest: One variable of interest (V22) was discarded from the analysis since the question was misinterpreted by a number of experts. Figure 1 represents the DM distributions for the first question (11 related variables of interest), related to the $\log_{10}$ increase or decrease in cfu's on a pig carcass. It can be read from the figure that the confidence bands (5% till 95% quantiles) are narrower for the item-weight DM, than for the other DMs. Figure 2 (variable of interest question 2) shows the DM distributions estimating the *Salmonella* prevalence in subsequent production stages at the abattoir. The confidence intervals in the item weight DM distributions, were similar to those obtained by the global weight DM, but were smaller than those of the equal weight and user weight DM, for most of the production processes

at the slaughterhouse (from singeing until chilling) indicating that combined distributions based on the performance based DMs were much more informative. In Fig. 3 (question of interest 3), it can be read that the confidence intervals for the four DMs were almost identical. The estimates of the 50% quantile were somewhat higher in the performance based DMs. The performance-based DMs produced higher estimates for the medians, than the equal weight and user weight DM (Fig. 4). In Fig. 5, almost no difference was observed between the item weight DM and the user and equal weight DM.

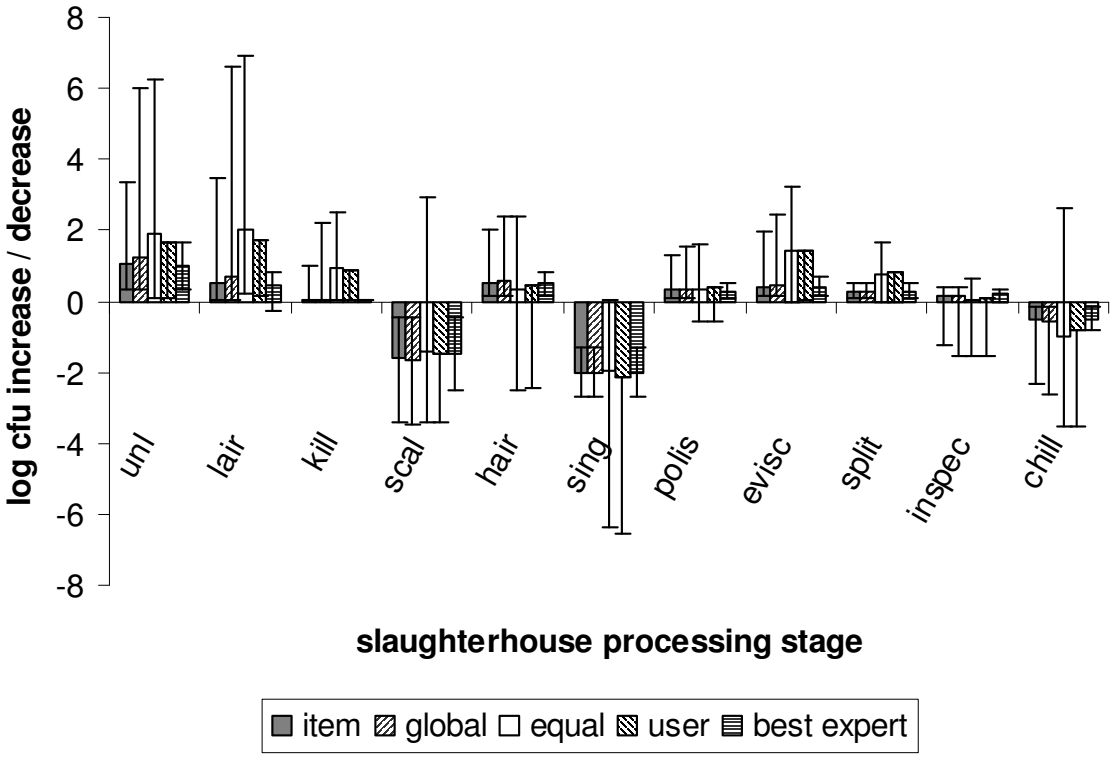


Fig. 1. The decision maker's distribution estimating the *Salmonella* concentration increase/decrease in $\log_{10}$ cfu at abattoir processing stages, expressed by the 5%, 50% and 95% quantiles, obtained by the item weight, the global weight, the equal weight, the user weight DM and the uncertainty distribution of the best expert. Unl=unloading pigs from truck slaughterhouse, lair=lairage, kill=stunning & killing, scal=scalding, hair=dehairing, singe=singeing, polis=polishing, evisc=evisceration, split=splitting, inspect=meat inspection, chill=chilling.

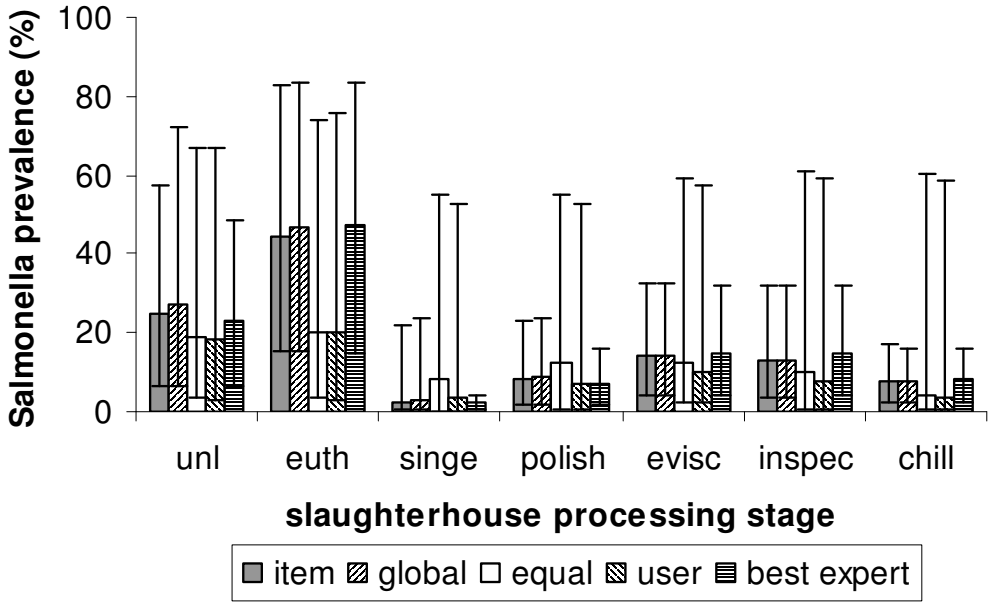


Fig. 2. The decision maker's distribution estimating the Salmonella prevalence at abattoir production stages, expressed by the 5%, 50% and 95% quantiles obtained by the item weight, the global weight, the equal weight, the user weight DM and the uncertainty distribution of the best expert. A starting bacteriological prevalence of 7% (90 CI 6-8%) was assumed when pigs were leaving the farm. Unl=unloading pigs from truck slaughterhouse, euth=euthanisia, sing=singeing, polis=polishing, evisc=evisceration, inspect=meat inspection, chill=chilling.

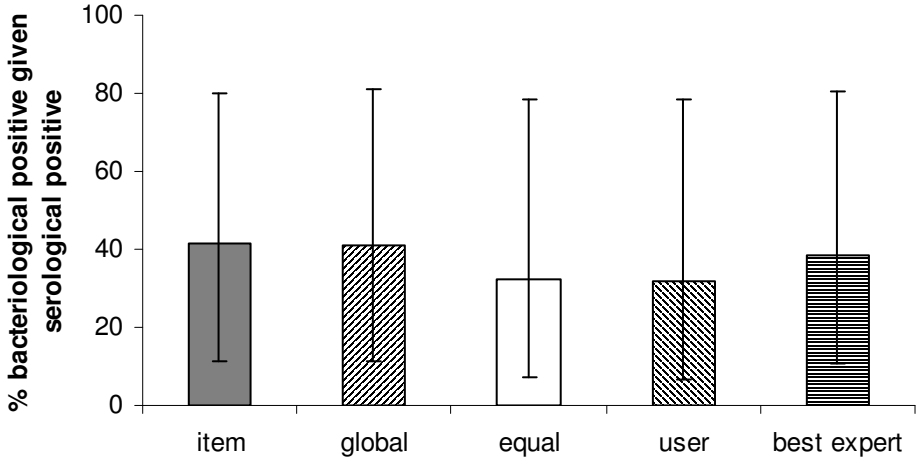


Fig. 3. The decision maker's distribution estimating the percentage of bacteriological Salmonella prevalence of a pig on the farm given it is also serologically positive, expressed by the 5%, 50% and 95% quantiles obtained by the item-weight, the global weight, the equal weight, the user weight DM and the uncertainty distribution of the best expert.

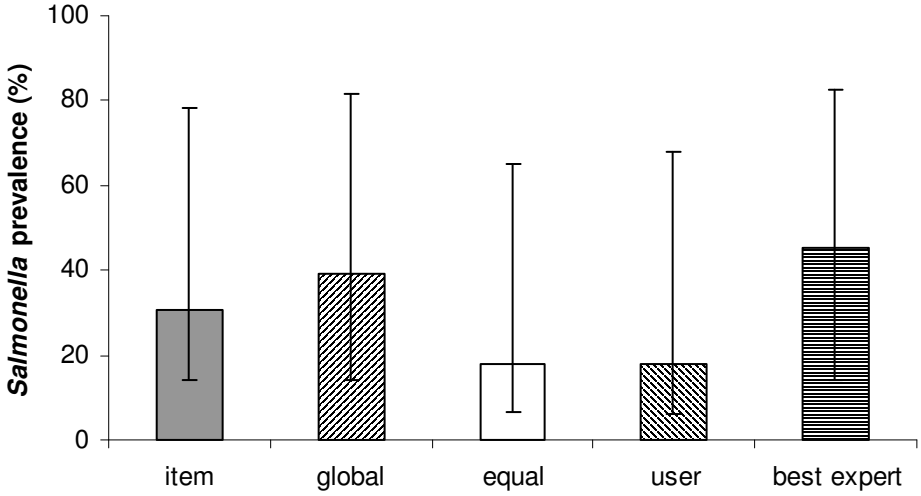


Fig. 4. The decision maker's distributions estimating the increase in bacteriological *Salmonella* prevalence (%) after transport and lairage expressed by the 5%, 50% and 95% quantiles obtained by the item weight, the global weight, the equal weight, the user weight DM and the uncertainty distribution of the best expert.

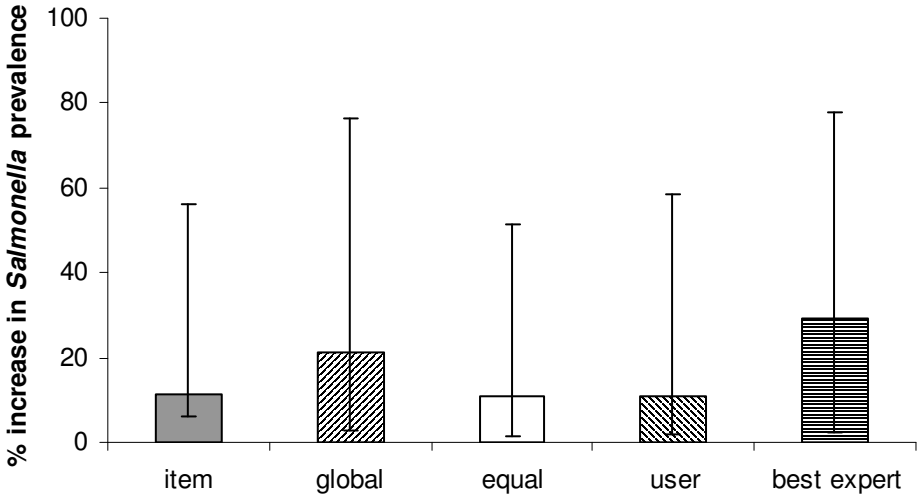


Fig 5. The decision maker's distributions estimating the increase in bacteriological Salmonella prevalence (%) due to improper cleaning of the conveyor belt and work surface at the cutting plant expressed by the 5%, 50% and 95% quantiles obtained by the item-weight, the global weight, the equal weight, the user weight DM and the uncertainty distribution of the best expert.

Experts' self-assessment of expertise: By attributing the experts' self-rating of expertise (sum of the experts' self-rating of expertise over all the seed variables) as user weights in the user-weight DM, the least calibrated and least informative DM was obtained. This suggests that in our study the experts' self-expertise is on average not indicative for their performance to the seed variables as measured by information and calibration, i.e. some experts were over or

underconfident. Experts 11, 9 and 4 were the experts that had the largest weights in the performance-based DM. Expert 11, who was the best experts according to his performance on the seed variables, also attributed him/herself the highest scores related to his/her self-expertise to the asked questions. Expert 10 was an example of an expert who was unweighted in the performance-based DMs, but attributed him/herself a weighting score as high as expert 11. Expert 4 was clearly an underconfident expert, with the 3rd largest unnormalized weight (0.01836), this expert attributed him/herself the lowest scores for his self-assessment of expertise.

DISCUSSION

In this study, a structured judgement approach was chosen in order to combine expert judgements into a combined distribution called the decision maker. The used protocol was different as compared to the one described by Cooke and Goossens (2000) and Van der Fels-Klerx et al. (2005), where experts are asked to present their PDFs by giving quantiles from the unknown distribution of the variables of interest and the seed variables, with for instance the 5%, 50% and 95% quantiles. The reason for eliciting expert judgement through minimum, most likely and maximum values has been chosen because it is a more straightforward way than to elicit estimates through quantiles, for which a specific training session is required (Cooke & Goossens, 2000). We aimed to enhance a rational consensus on the quality of these combined distribution, by following a structured and traceable procedure so that the expert judgements can be used as scientific data. Different weighting schemes for combining expert subjective probability distributions serving as input for a QMRA-model were compared.

According to the experts' performance on the seed variables, the item-weight DM obtained the highest calibration and information score, as compared to the global weight, equal weight and user weight DM. This item weight DM outperformed the best expert in terms of his unnormalized weight (virtual weight). There is no mathematical theorem that the performance-based decision maker outperforms the equal weight DM, but in practice the performance-based DM is usually better than the equal-weight DM (Cooke & Goossens, 2006). The use of a performance-based DM can be motivated by the fact that the experts' judgements were elicited from a heterogeneous panel of experts originating from nine different countries and with different backgrounds, as was also highlighted by Van der Fels-Klerx et al. (2002). The combined PDF for the variables of interest obtained from the item weight DM can be readily be used to provide the input parameters for the QMRA, since this DM obtained the highest performance. The performance-based DM was more informative than the equal weight and user weight DM.

The success of the implementation of the classical model depends to a large extend on finding the adequate seed variables. Indeed the performance of the experts to the seed variables is judged as indicative for their performance on the variable of interest. Concerning the minimal number of required seed variable, the number of effective seed variables (11) in the present study was judged successful. Goossens and Cooke (2005) state that the more seed variables the better, but that ten seed variables should be sufficient. Using seed variables offers an objective method to calibrate experts, but the calibration may not be correct in case inappropriate seed variables were presented to the experts. The choice for finding the ideal set of seed variables was difficult in this study, but considered adequate. Seed variables must resemble as much as possible the variables of interest, and the realisation of the seed variables must be readily accessible to the analyst during the time of the study. In this study seven seed variables were

discarded due to e.g. prior knowledge of an expert and/or missing values. Many experts argued that providing the estimates for the seed question and variables of interest was (very) difficult. The user-weight DM which was given weight according to the experts' self-assessment of expertise was judged unsatisfactory. A high (low) self-rating of expertise did not result in a high (low) DM virtual weight. Although the user-weight DM was better calibrated than the best expert while its information score was the lowest of all the evaluated decision makers. The virtual weight of the item weight DM was lower than the equal weight DM. We conclude that the use of the self-rating of experts did not provide a rational objective basis in this study for weighting the experts. The performance on seed variables offers a more objective basis for weighting the experts.

The used expert judgement protocol used to combine PDFs for the variables of interest can be used in the QMRA model for *Salmonella* in pigs. The performance-based DM was useful since it yielded more informative distributions than the other weighting schemes. The proposed protocol is judged useful to evaluate weighting schemes and combine PDFs to provide input in future QMRA models, and is likely to enhance the transparency of the QMRA process.

ACKNOWLEDGEMENTS

This study has been carried out with the financial support of the Belgian Federal Public Service of Health, Food Chain Safety and Environment research programme (R-04/003-Metzoon) 'Development of a methodology for quantitative assessment of zoonotic risks in Belgium applied to the '*Salmonella* in pork' model'. Partners in the METZOON research consortium are the Veterinary and Agrochemical Research Centre (VAR), the Schools of Veterinary Medicine of both Liège and Ghent University, the Institute for Agricultural and Fisheries Research (ILVO), the federal Institute for Public Health and the Center for Statistics of Hasselt University.

We kindly thank all the experts who participated in this study for their collaboration. We gratefully acknowledge Dr. Dorota Kurowicka and Prof. Roger Cooke (Delft University of Technology, Delft, the Netherlands) for the constructve discussions.

REFERENCES

Alban, L. and Stark, K.D. (2005). Where should the effort be put to reduce the Salmonella prevalence in the slaughtered swine carcass effectively? Prev. Vet. Med. 68, 63-79

Clemen, R.T. and Winkler, R.L. (1999). Combining Probability Distributions From Experts in Risk Analysis. Risk Anal. 19, 187-203

Cooke, R.M. (1991). Experts in Uncertainty. Opinion and Subjective Probability in Science. Oxford University Press, New York

Cooke, R.M. and Goossens, L.H.J. (2000). Procedures Guide for Structured Expert Judgement. (Brussels, Belgium, European Commission, Directorate-General for Research), Technical report EUR 11820 EN

Cooke, R.M. and Goossens, L.H. (2006). TU Delft Expert Judgment Data Base (Washington, Expert Judgment Policy Symposium and Technical Workshop), 25p, available at: http://dutiosc.twi.tudelft.nl/~risk/

Cooke, R.M. and Goossens, L.L.H.J. (2008). TU Delft expert judgment data base. Reliability Engineering & System Safety. 93 (5), 657-674

EFSA (2006). The Community Summary Report on trends and sources of zoonoses, zoonotic agents, antimicrobial agents, antimicrobial resistance and foodborne outbreaks in the European Union in 2005. The EFSA Journal 94

Goossens, L.H. and Cooke, R.M. (2005). Expert Judgement - calibration and combination. In Workshop on expert judgment (Aix En Provence, France, June 2005), available at: http://dutiosc.twi.tudelft.nl/~risk/

Lopez De La Cruz, J. (2004). Application of probability models and expert judgement analysis in information security. Department of Applied Mathematics and Risk Analysis, Delft University of Technology, M.Sc. thesis

Nauta, M.J., Evers, E.G., Takumi, K. and Havelaar, A.H. (2001). Risk assessment of Shiga-toxin producing Escherichia coli O157 in steak tartare in the Netherlands (Bilthoven, RIVM), p. 169p

Scholz, R.W. and Hansmann, R. (2007). Combining Experts' Risk Judgments on Technology Performance of Phytoremediation: Self-Confidence Ratings, Averaging Procedures, and Formative Consensus Building. Risk Anal. 27, 225-240

Stark, K.D., Wingstrand, A., Dahl, J., Mogelmose, V. and Lo Fo Wong, D.M. (2002). Differences and similarities among experts' opinions on Salmonella enterica dynamics in swine pre-harvest. Prev. Vet. Med. 53, 7-20

Van der Fels-Klerx, H.J., Cooke, R.M., Nauta, M.N., Goossens, L.H. and Havelaar, A.H. (2005). A structured expert judgment study for a model of Campylobacter transmission during broiler-chicken processing. Risk Anal. 25, 109-124

Van der Fels-Klerx, I.H., Goossens, L.H., Saatkamp, H.W. and Horst, S.H. (2002). Elicitation of quantitative data from a heterogeneous expert panel: formal process and application in animal health. Risk Anal. 22, 67-81

Van der Gaag, M. and Huirne, R.B.M. (2002). Elicitation of expert knowledge on controlling Salmonella in the pork chain. Chain and network science, 135-147

Vargas Galindo, A.P. (2007). Probabilistic Inversion in Priority Setting of Food Borne Pathogens. Department of Applied Mathematics and Risk Analysis, Delft University of Technology, M.Sc. thesis

Vose, D. (2000). Risk analysis - A quantitative guide, 2nd edition Edition. John Wiley and Sons publishers Ltd., Chichester

Walker, K.D., Catalano, P., Hammitt, J.K. and Evans, J.S. (2003). Use of expert judgment in exposure assessment: part 2. Calibration of expert judgments about personal exposures to benzene. J. Expo. Anal. Environ. Epidemiol. 13, 1-16

CATTLE DISEASES

MASTITIS INCIDENCE EXPLAINED BY FARMERS' ATTITUDE AND BEHAVIOUR

J. JANSEN*, B.H.P. VAN DEN BORNE, R.J. RENES, G. VAN SCHAIK, T.J.G.M., LAM, AND C. LEEUWIS.

SUMMARY

Several studies suggest that farmers' attitudes and behaviour towards different aspects of mastitis can explain the variation in mastitis incidence on farms. A survey on self reported attitudes, behaviour and mastitis incidence was conducted on 336 Dutch dairy farms to determine and to quantify the added value of farmers' attitudes, beyond the farmers' behaviour, in explaining the variation of mastitis incidence. Results of two-step stepwise multiple linear regression analyses show that farmers' self-reported behaviour and attitudes together explained 48%, 29% and 23% of the variation of the average farm bulk milk somatic cell count, the clinical mastitis incidence and the combined (sub)clinical mastitis incidence, respectively. This variance was mostly explained solely by the attitude variables. The results of this study show that farmers' attitudes were a better measure to explain and predict differences in mastitis incidence between farms than farmers' self reported behaviour.

INTRODUCTION

An increase in mastitis incidence is due to either increased infection pressure or decreased cow resistance. The latter can be caused by factors outside the farm (such as weather), but it usually indicates that farm management is not optimal. Numerous quantitative studies have demonstrated the effect of farm management practices on mastitis (Elbers et al., 1998; Barkema et al., 1999; Barnouin et al., 2005; Chassagne et al., 2005; Green et al., 2007; Nyman et al., 2007; Wenz et al., 2007). However, in these studies the identified risk factors could only explain a part of the variance of mastitis incidence on farms. Preventive as well as treatment programs, based on known risk factors of mastitis, sometimes fail for reasons that are not immediately understood by the health professionals connected to the dairy herd (Vaarst et al., 2002). Several studies suggest that whether and how these mastitis management practices are implemented on a farm probably depends on the "human factor" of the farmer, including his management style and accompanying dispositions and beliefs (i.e. 'attitudes') towards different aspects of mastitis treatment and preventive behaviour (Dohoo et al., 1984; Seabrook, 1984; Tarabla and Dodd, 1990; Beaudeau et al., 1996; Barkema et al., 1999; Reneau, 2002; Vaarst et al., 2002; Andersen and Enevoldsen, 2004; Barnouin et al., 2004; Leeuwis, 2004; Nyman et al., 2007; Wenz et al., 2007).

In the field of social sciences, the impact of the "human factor" on behaviour is widely studied by using constructs like peoples' attitudes, knowledge, beliefs, values, goals and intentions (Jaccard and Blanton, 2005). Attitude is especially well known as an important factor

* Jolanda Jansen, Wageningen University, Communication Science, Wageningen, The Netherlands, P.O. Box 8130 6700 EW, email: Jolanda.Jansen@wur.nl

in (changing) behavioural intentions and actions (Ajzen and Fishbein, 2005). The term attitude is used for evaluative tendencies, which can both be inferred from and have influences on cognitive beliefs, affective associations and overt behaviour (Albarracin et al., 2005; Kiviniemi et al., 2007). In this study, the construct of attitude is used as a collective term for these cognitive beliefs and affective associations in which issues like knowledge, beliefs, values, goals and intentions are included.

Although many studies in agricultural sciences indirectly implicate attitude as a determining risk factor for mastitis incidence, there have been few studies that have attempted to directly correlate farmers attitude with milk quality (Reneau, 2001, 2002). Preliminary research of the Dutch Udder Health Centre (Van der Zwaag et al., 2005) propose that farmers attitude might indeed be more correlated to mastitis incidence than farmers behaviour (Kuiper et al., 2005). Moreover, exploratory research of Tarabla and Dodd (1990) and Bigras-Poulin et al. (1985) showed that farmers' attitudes, values and socio-demographic profile explained a similar or greater amount of the variation in some farm performance characteristics than just farm management variables. Unfortunately, these studies fail to explain which attitude is important and how this specifically relates to incidence of diseases such as mastitis. Although many studies already suggest an important effect of attitude on farm performance, the direct effect of farmers' attitude on (sub)clinical mastitis incidence is, so far, hardly investigated. In this study the variance of different mastitis incidence indicators between farms is explained by both behavioural items as well as attitudinal items by answering three questions:

- First, is it possible to explain mastitis incidence by using self reported behaviour and attitude of farmers?
- Second, does the farmers' attitude have a quantifiable added value beyond the farmers' behaviour, in explaining the variation of mastitis incidence?
- Third, which specific behavioural and/or attitudinal variables are then most important in explaining this mastitis variance?

The answers to these questions will contribute to the understanding of mastitis problems and to provide leads for effective communication strategies in mastitis control programs.

MATERIALS AND METHODS

General

An extensive survey was carried out as part of the Dutch udder health program. The independent variables about both attitude and behaviour were obtained with a questionnaire. The dependent variables of mastitis incidence indicators were observed by the farmer or measured via test day records and bulk milk somatic cell count (BMSCC) data.

Criteria for farms to participate in the study were: (1) an average farm size > fifty cows, (2) the age of the farmer had to be < 57 years, and (3) farms had to participate in the regular test-day recording with test day intervals of 3-6 weeks. These criteria were used to ensure that these farms would be able to participate in the Dutch udder health program for the coming years. The selection resulted in a random sample of which 543 farmers were contacted by telephone to ask for cooperation. Subsequently, 378 participants completed a questionnaire on attitudes and behaviour and gave their permission to collect mastitis data. Farmers who did not want to cooperate with the survey mentioned that they were either too busy or not interested. After one year of data collection, complete records of 336 farms were available for analysis, which was

equal to a total response rate of 62%. Reasons for missing data in the survey were (1) farmers who quit farming or reorganized the farm, (2) farmers who had neither time nor willingness to fill in the papers and (3) farmers who provided incomplete data, such as incomplete or missing forms.

Mastitis data as dependent variables

Three different mastitis indicators were used in this study as the dependent variables to provide insight in the mastitis status of the farms: (1) clinical mastitis incidence rate, (2) average BMSCC in the period preceding the survey (April 2004 - July 2004) and (3) a combined (sub)clinical mastitis incidence rate.

The clinical mastitis incidence rate was calculated as the number of clinical cases divided by the number of days at risk. Clinical mastitis cases were reported by the farmer by describing cows' registration number as well as the date and the quarter of the infection. A clinical mastitis case was defined as a cow with visual abnormalities in the milk and/or quarter. Clinical mastitis cases occurring within two weeks from each other in the same quarter were excluded from the analysis. Cow days at risk were calculated as the total number of days a cow was present at the farm during the study. The clinical mastitis incidence rate was then calculated at herd level using the following equation, expressed as the incidence rate per 100 cows at risk per year per farm:

$$\text{Clinical mastitis incidence rate} = (\text{new}/\text{dar}) * 365 * 100 \quad (1)$$

when new = number of new cases of clinical mastitis a year and dar = number of days at risk for clinical mastitis

For the second mastitis indicator, the fortnightly BMSCC data were used to calculate the average BMSCC of the three months preceding the survey. The BMSCC data preceding the survey was considered most appropriate because the questionnaire covered this period.

For the third mastitis indicator, the combined (sub)clinical mastitis incidence rate, farm data of both clinical and subclinical mastitis were analyzed. The following equation was used to calculate the combined (sub)clinical at herd level, expressed as an incidence rate per 100 cows at risk per year per farm:

$$\text{(Sub)clinical mastitis incidence} = ((\text{newclin} + \text{newsubclin}) / \text{darmast}) * 365 * 100 \quad (2)$$

when newclin = number of new cases of clinical mastitis a year, wewsubclin = number of new cases of subclinical mastitis a year, darmast = number of days at risk for clinical and/or subclinical mastitis

In this combined measure, a new case of mastitis was considered to be a new case of clinical mastitis as described above, or as a new case of subclinical mastitis, based on composite somatic cell counts (CSCC), gathered from the regular test day recording. A new case of subclinical mastitis was defined as an increase above a threshold value of 150,000 cells/ml for heifers and 250,000 cells/ml for multiparous cows, after being two consecutive test days below these threshold values, regardless the dry period. These threshold values are generally used in the Netherlands. As such, the cows could get more than one (sub)clinical mastitis case in the same

lactation. The (sub)clinical mastitis data was also corrected for cow days at risk. A cow was defined to be at risk if it had a low SCC and no clinical mastitis.

Attitudinal and behavioural data as independent variables

The data were collected by a structured questionnaire on self reported attitudes and behaviour as well as demographic items. A team of veterinarians, farmers, animal health experts, communication experts and social psychologists developed this questionnaire which contained 55 items regarding behaviour and 123 items about attitude. Insights about self-reported behaviour of farmers regarding mastitis were obtained by asking about actual actions, such as 'do you clean the teats before milking' or 'how often do you clean the cubicles'. Insights about self reported attitude were mainly derived by asking about perceptions and opinions such as 'I worry about mastitis" and 'what is the most annoying aspect of mastitis'.

The attitude and behaviour items were mostly measured by statements that the farmers rated on a five-point Likert Scale (Likert and Hayes, 1961) according to how much they agreed or disagreed with the statements, for example: 'I worry about mastitis (1= completely disagree to 5=completely agree)'. In addition, binominal items were used to see whether they answered yes (1) or no (0) on a certain question, e.g. 'do you disinfect all teats after milking with a spray or dip'. Finally, some items were measured continuously, e.g. age and BMSCC level, or by categorizations in groups, e.g. type of milking parlour. Farmers' normative frame of reference was measured by asking farmers when, at which value, they perceive a problem with BMSCC and clinical mastitis, and when they are satisfied. All the items from the questionnaire were used to develop a set of independent variables by Principal Component Analysis (PCA) to explain the variation in mastitis incidence.

Data analysis

The data analysis is based on three steps: (1) reducing the number of variables from the questionnaire by PCA, (2) correlation analysis with this reduced set to determine associations with mastitis indicators, (3) regression analysis with the variables that correlated significantly with at least one of the mastitis indicators.

For the first step, PCA with Varimax rotations and reliability analyses were performed on items which were measured on the same Likert Scale (Dohoo et al., 1997; Field, 2005). Factors with an eigenvalue >1 (Kaiser criteria (Kaiser, 1960; Dohoo et al., 1997)) were included and tested for reliability using Cronbach's $\alpha>.55$ as threshold value for combining items in the same measure. These new multiple item measures were computed for all farmers by taking the average score of the underlying variables. The multiple item measures were used in further analyses. Items which could not be grouped based on PCA and reliability were regarded as independent variables and were included individually in the analyses. As presented in Appendix A and B, the PCA and reliability analyses resulted in 46 behavioural variables including two multiple-item measures, and 95 attitudinal variables including twelve multiple-item measures.

For the second step, after the PCA, the data were split in two main parts for the analyses: (1) 46 self reported behavioural variables from the questionnaire and (2) 95 self reported attitude variables from the questionnaire. In these analyses, the three (sub)clinical mastitis indicators were used as dependent variables and the self-reported measures of attitudes and behaviour as independent variables. To select the variables from both behavioural and attitudinal variables with a significant ($p<.05$) association with one of the dependent variables, zero-order bi-variate

two tailed Pearson correlations were calculated. After correlation, the variables were grouped on subject, e.g. 'milking procedures' or 'perception of control'. The significant variables were used in the final step to test whether and how much farmers' attitudes as well as their behaviour explained the between-farm variation of mastitis incidence.

In the final step, linear multiple two-step stepwise regression analyses were performed. All attitudinal and behavioural variables that were used in the regression analyses correlated significantly ($p<.05$) with at least one of the mastitis indicators (Dohoo et al., 1997). In the first model, only behavioural variables, in the second model only attitudinal variables and in the third model all variables were included. The model was checked for normality and autocorrelation. Data were analyzed using SPSS (SPSS 12.0.1 for Windows, SPSS Inc., Chicago, Illinois, U.S.A.).

RESULTS

General description of farms

An exploration of the mastitis indicators showed that the average clinical mastitis incidence was 30.3 cases per 100 cows at risk per farm per year (SD 17.70). Furthermore, the average BMSCC preceding the survey was 191,890 cells/ml (SD 61.04) and the average (sub)clinical mastitis incidence was 99 cases per 100 cows at risk per farm per year (SD 29.83).

The average herd size of the participants was 77 dairy cows (SD 23.86). Less than two percent of the respondents had an organic farming system. The farmers were on average 41 years old (SD 8.35) and almost all farmers had followed secondary education of whom 68% followed intermediate professional agricultural education and 13% followed higher professional agricultural education. Milking systems like automatic milking systems or carousel milking parlours were used by one and three percent of the farmers, respectively.

Explaining the variance of mastitis incidence

As presented in Table 1, 2 and 3, the results of the two-step stepwise multiple linear regression analyses show that farmers' self-reported behaviour and attitudes together explain 48%, 29%, and 23% of the variation within the average farm BMSCC, the clinical mastitis incidence and the combined (sub)clinical mastitis incidence rate, respectively. The variables from the final set of independent data did not correlated higher than .34 with each other. In addition, Variance Inflation Factor (VIF) values were reported around 1.00 and Durbin Watson-tests were reported around 2.00, which indicate independent errors and low multicolinearity. Furthermore, analyses of standardized residuals were normally distributed and gave no reasons for concern.

Although behavioural variables, as well as attitudinal variables, were able to predict unique variance in all mastitis indicators, the results show that the variance in mastitis incidence is mainly explained by farmers' attitudes. As presented in Table 1, 2 and 3, 47% of the variance in BMSCC, 27% of the variance in the clinical mastitis incidence rate and 17% of the variance in the combined (sub)clinical mastitis incidence rate was explained by just the attitude variables.

The best explaining variables for mastitis incidence

The stepwise regression method showed that the variation in clinical mastitis was best explained by the perception of control of mastitis (β=.38, p<.001). When explaining clinical mastitis incidence rate, attitudes played a significant role in the final model. Table 1 shows that the only significant behavioural variable which turns out to be significant in the final model is "to account for udder health parameters when selecting bulls".

Table 1. Clinical mastitis incidence[a] explained by attitude and behaviour of Dutch Dairy farmers with two step stepwise multiple linear regression analyses.

MODEL [b]	1	2	3
Behavioural variables			
Milking procedures			
Cleaning udders before milking with paper towel	.16*		.09
Fore strip all cows before milking	.14*		
Treatment and actions			
No actions are taken as long as there are no serious mastitis problems	-.19**		-.11
Other			
Accounting for udder health parameters when selecting bulls for mating	.17**		.19**
Attitudinal variables			
Normative frame of reference			
Satisfaction level percentage clinical mastitis cases		.18**	.19**
Perceived effect of penalty level			
The best way to decrease BMSCC nationally is to decrease the penalty level		.12*	
Perception of control			
Perceived lack of control of mastitis		.37***	.38***
Bad luck plays an important role in a mastitis outbreak		.18**	.18**
Worrying about mastitis		.15*	
Knowledge			
Perceived knowledge of mastitis treatment		.20**	.14*
Other			
The best way to decrease BMSCC nationally is free visit of mastitis expert every month		-.13*	-.16**
High milk production per cow is important farm goal		.19**	.18**
Too little time to work on mastitis prevention		-.19**	-.19**
Model *F*	7.19***	10.83***	10.69***
Df	(4,236)	(9,231)	(10,230)
R^2	.11	.30	.32
Adjusted R^2	.09	.27	.29

[a] Incidence rate of clinical mastitis cases per 100 cows per year

[b] Coefficients are standardized regression weights (betas). *p<.05**p<.01***p<.001 . Exclude cases listwise Only those variables are presented that were significant in at least one model of the stepwise multiple linear regression analyses.

The variation in BMSCC value was best explained by (1) the farmers' normative frame of reference about mastitis (β=.33, p<.001), (2) the perception of control of mastitis (β=.25, p<.001), and (3) the perceived effect of the BMSCC penalty level (β=.24, p<.001). When explaining the variance in BMSCC attitudinal variables also seemed to be most important. Table 2 shows that the only behavioural variable which turns out to be significant in the final model is "to check cell count records of individual cows although BMSCC is low".

Table 2. Bulk Milk Somatic Cell Count (BMSCC)[a] explained by attitude and behaviour of Dutch Dairy farmers with two step stepwise multiple linear regression analyses.

MODEL [b]	1	2	3
Behavioural variables			
Milking procedures			
Wearing gloves during milking	-.16**		
Treatment and actions			
Individual cows' cell count records are not checked when BMSCC is low.	.20**		.14**
Strictly finish antibiotic treatment	-.13*		
Other			
Frequency of cleaning cubicles every day	-.19**		
Attitudinal variables			
Normative frame of reference			
Satisfaction level percentage subclinical mastitis cases		.11*	
Perceived frame of reference BMSCC [c]		.30***	.33***
Perceived effect of penalty level			
Perceived effect on farmers behaviour if BMSCC penalty level decreases from 400.000 to 350 000 cells/ml		.26***	.24***
No change in clinical mastitis treatment when penalty level decreases to a BMSCC of 350.000 cells/ml		-.15**	-.15**
The best way to decrease BMSCC nationally is to decrease the penalty level		.11*	.11*
Perception of control			
Perceived lack of control of mastitis		.24***	.25***
Knowledge			
Perceived knowledge of mastitis treatment			
Other			
The best way to decrease BMSCC nationally is free visit of mastitis expert every month		.18***	.18***
Model *F*	9.15***	31.16***	32.04***
Df	(4,236)	(7,33)	(7,233)
R^2	.13	.48	.49
Adjusted R^2	.12	.47	.48

[a] Average fortnightly BMSCC of three months preceding the survey.
[b] Coefficients are standardized regression weights (betas). *p<.05**p<.01***p<.001. Exclude cases listwise. Only those variables are presented that were significant in at least one model of the stepwise multiple linear regression analyses.
[c] The average of the perceived problem level of BMSCC and the perceived satisfaction level of BMSCC

The variation in the combined (sub)clinical mastitis incidence rate was best explained by the perceived effect of a BMSCC penalty level (β=.25, p<.001) and the frequency of contact with others (β=.24,p<.001). When explaining the combined (sub)clinical mastitis incidence rate of a farm, model 3 of table 3 shows that this mastitis indicator has more significant behavioural variables related to mastitis than the other two mastitis indicators.

Table 3. A combined measure of (sub)clinical mastitis incidence[a] explained by attitude and behaviour of Dutch Dairy farmers with two step stepwise multiple linear regression analyses.

MODEL [b]	1	2	3
Behavioural variables			
Treatment and actions			
Delayed treatment of subclinical mastitis cows when milk quota is not full	.12*		
Percentage of mastitis cases from which milk samples are taken for bacteriology	-.19**		-.14*
Other			
Frequent contact with others about mastitis	.26***		.24***
Frequency of cleaning cubicles every day	-.22***		-.16**
Attitudinal variables			
Perceived effect of penalty level			
Perceived effect on farmers behaviour if BMSCC penalty level decreases		.29***	.25***
Perception of control			
Perceived lack of control of mastitis		.15*	.16*
Most annoying aspect of mastitis is uncertainty if cow recovers		-.13*	
Had a serious mastitis problem once		.12*	
Model *F*	10.34***	13.51***	14.94***
Df	(4,236)	(4,236)	(5,235)
R^2	.15	.19	.24
Adjusted R^2	.14	.17	.23

[a] A combination of clinical mastitis incidence and subclinical mastitis incidence per 100 cows per year
[b] Coefficients are standardized regression weights (betas). *p<.05**p<.01***p<.001 . Exclude cases listwise . Only those variables are presented that were significant in at least one model of the stepwise multiple linear regression analyses.

DISCUSSION

Explaining mastitis incidence by self reported behaviour and attitudes

The aim of this study was to explain the variance in mastitis incidence between farms by both behavioural items as well as attitude items of Dutch dairy farmers. The results suggest that indeed mastitis can be explained to a certain extend by farmers attitudes and behaviour and that attitudes have a quantifiable added value in these models. In our study, farmers' attitudes explain 19% to 48% of the variance in mastitis indicators, while farmers' self reported behaviour explains 9% to 14% of the variance. An early study of Bigras-Poulin et al. (1985) showed the effect of attitudes on farm performance. They found that socio-psychological variables explained 11% to 25% of the variation, while management variables explained 0% to 16% of the variation in reproductive performances of the herd. These results support our findings that

attitudes should be taken into account when studying farm performances. In addition, research of Tarabla and Dodd (1990) showed that in most of their models the variables related to farmers attitudes explained a similar or greater amount (between 14% and 35%) of the variation of farm performance than the group of management variables (between 14% and 26%). Although their study design was slightly different, clinical mastitis was not included, it was concluded that human variables could explain why there is still a large variation in milk quality and milk production among farmers after years of improvements in the dairy sector.

In social psychology, attitudes are regarded as one of the main factors influencing behaviour (Ajzen and Fishbein, 2005). Although in this study attitudes and behaviour are separated, it needs to be considered that the attitude influences farmers' behaviour and therefore can influence the health status of a farm. The results of this study strengthen this belief, because the self reported behaviour of farmers did not seem to influence mastitis incidence to a great extent. This self reported behaviour is probably influenced by farmers' attitudes, e.g. if you ask farmers if they clean the udders before milking, one still do not know how clean these udders will be. This study therefore suggests that self reported behaviour insufficiently explains farm management and farm performances.

Interestingly, mainly attitudes regarding farmers' frame of reference and perception of control have strong associations with mastitis incidence. Regarding the farmers' normative frame of reference this study indicated that what a farmers regards as a serious mastitis problem differed among farmers. These normative beliefs imply an action moment. Only when farmers regard the mastitis incidence on their farms as problematic they will take actions. Regarding the farmers' perception of control, research showed that a lack of feeling of control could endow their capacity to deal with the real situation. As such, mastitis control becomes a self–fulfilling prophecy, as long as farmers do not believe they can control the situation, they will not feel able to take (preventive) measures, and consequently will have more problems.

The difference between different forms of mastitis

This study showed that the different forms of mastitis seemed associated with different attitudes and behaviour of farmers. The results suggest that apparently it is easier to explain BMSCC values than mastitis incidence by self reported attitude and behaviour. An explanation could be that the "farmer factor" in addition to the "cow factor" and the "pathogen factor" is more important in BMSCC control than in clinical mastitis control, possibly because BMSCC levels can more easily be managed (e.g. by excluding high SCC milk from the tank or by culling cows with (sub)clinical mastitis). In addition, the results showed that the variance in clinical mastitis problems was explained by different attitudinal and behavioural items than the variance in BMSCC. On the one hand, the regression analyses showed that farmers with clinical mastitis seem to arrange time to work on mastitis prevention and account for udder health parameters when selecting bulls. On the other hand, the analyses showed that the BMSCC level is mainly explained by the normative frame of reference and the perceived effect of a penalty level, which can be regarded as external motivators. This might suggests that farmers' with clinical mastitis problems are more aware of mastitis compared to farmers' with BMSCC problems. This could be due to the fact that clinical mastitis is a direct visible problem, while BMSCC problems can only be identified when checking cell count records.

Mastitis is not only a technical issue

From a historical perspective, agricultural extensionists, researchers and veterinarians assumed that agriculture was a separate activity executed by an individual farmer that was based primarily on rational, technical and economic considerations (Leeuwis, 2004). Although, these rational choices still play a role in farm management, we have learned that farmers' decision making based on these considerations is not always clear and understandable (Vaarst et al., 2002). Nowadays, many studies suggest the effect of the "human" factor on farm performance (Willock et al., 1999; Bergevoet et al., 2004; Leeuwis, 2004). Barkema et al. (1999) studied management styles and the associations with BMSCC and clinical mastitis. Their study showed that farmers that were regarded as clean and accurate were associated with lower BMSCC levels, while farmers regarded as quick and dirty were associated with higher BMSCC levels. They concluded that management did have an influence on the implementations of measures to prevent mastitis (Barkema et al., 1999). In addition, several other studies suggested the effect of attitudes, such as awareness of the farmer on mastitis (Hutton et al., 1990; Chassagne et al., 2005; Nyman et al., 2007; Wenz et al., 2007). Although many studies implied the importance of farmers' attitudes on farm performance, there have been few studies that have attempted to directly correlate farmers' attitude with milk quality. As far as known only exploratory research of Tarabla and Dodd (1990) and Bigras-Poulin et al. (1985) showed quantifiable effect of socio-psychological variables on farm performance regarding reproductivity and milk production. As such, to the authors' knowledge, this study seems to be the only recent, empirical investigation showing that farmers' attitudes have a significant quantifiable effect in explaining mastitis incidence on farms in addition to farmers' behaviour.

Some limitations of the study

Despite the fact that the results of this study are supported by findings in literature, this study has its limitations. The studied population was a random sample from younger farmers with larger herds. This was specifically done to include all farmers who were expected not to stop farming in the coming years and therefore could contribute to milk quality in the Netherlands. Additionally, all farms participated in CSCC recording every 3-6 weeks, while other intervals or not testing for CSCC also exists in the Netherlands. The results of this study might therefore not apply to the whole Dutch dairy sector. In addition, the farmers participating in this study could be different than the average Dutch farmers, because they were willing to participate (selection bias).

It is important to note that this study was based on self reported attitudes and behaviour of farmers. It is possible that social desirable answers were reported by the farmers, which could have lead to a bias in the results. It is also important to note that although the survey was extensive and developed with mastitis experts, the total dataset of farmers' attitudes and behaviour regarding mastitis could be incomplete, which could explain why the survey was not able to explain more than 50% of the variance in mastitis incidence. Furthermore, it should also be taken into account that although in this study farmers' behaviour and attitudes were presented as independent variables, self reported behaviour and attitudes can be related.

Another critical note concerns the collection of mastitis data, noting that clinical mastitis was defined as a cow with abnormal milk and/or udder. Farmers might have interpreted this definition differently and this could have resulted in an overestimation or underestimation of the incidence rate of clinical mastitis. Regarding the use of BMSCC as an indicator for subclinical mastitis on a farm, it has been suggested that other data, like the arithmetic average test day

somatic cell counts of the herd (HSCC), might be a better parameter (Lievaart et al., 2007). As such, it can be assumed that in this study, the effect of farmers' attitude and behaviour is underestimated, because the "real" HSCC level is supposed to be higher than the BMSCC because of the influence of individual cow yield and farmers withholding the milk from high SCC cows.

Finally, the results of this study indicate observed relationships, but the causality of these relationships is difficult to determine. However, despite these limitations, the results are consistent with findings in literature and provide insight in farmers' behaviour and attitudes and the effect on mastitis incidence.

Applications for mastitis control programs and future research

Currently, most mastitis control programs are focused on influencing farmers' behaviour. Taking into account the results of this study, more attention for farmers' attitude is needed to design effective (mastitis) control programs in the future. The application and actual prevention of dairy health problems requires understanding of the farm as an integrated system and most of all requires education and motivation of the farmers to implement the right management practices (Chase et al., 2006; LeBlanc et al., 2006). In communications with farmers, more efforts should be made to improve farmers' normative frame of reference and to improve farmers feeling of control of the mastitis situation. Moreover, mastitis control programs should make distinctions in the forms of mastitis. Farmers with clinical mastitis problems, farmers with high BMSCC levels, and farmers with both problems need to be addressed differently because different attitudes and behaviour play a role.

Future epidemiological studies on farms should not only take farmers behaviour into account when explaining the difference between farms, because management style can confound the relationship between actual risk factors and disease incidence (Barkema et al., 1999). Moreover, this study shows that self reported behaviour hardly explains mastitis incidence. An often used alternative is to observe farmers' behaviour. However, it is still difficult to describe farmers' real behaviour, because the observant might influence the farmer. Based on the results of this study, it can be suggested that measuring farmers' attitude might be a good alternative when studying risk factors of mastitis.

CONCLUSION

This study showed that farmers' attitudes and self-reported behaviour explained the variation in mastitis incidence to a certain extent. The results indicated that farmers' self reported behaviour explained the variation in mastitis incidence to a limited extend. Farmers' attitudes explained a significant larger part of the variation in mastitis incidence. Especially, the perceived feeling of control, the perceived effect of BMSCC penalty level and the normative frame of reference are important in explaining the variation in mastitis incidence. Furthermore, the results suggest that BMSCC levels are better explained by attitudes and self reported behaviour than (sub)clinical mastitis incidence. The results showed that clinical mastitis incidence was associated with different attitudes and behaviour than BMSCC. It can be concluded that farmers' attitudes are a better measure to explain differences in mastitis incidence between farms than farmers' self reported behaviour and should therefore be taken into account in future research and animal health promotion.

ACKNOWLEDGEMENTS

This study is part of the five year mastitis program of the Dutch Udder Health Centre and was financially supported by the Dutch Dairy Board. The authors gratefully acknowledge the efforts from the farmers who participated in this study. The Dutch Cattle Syndicate (NRS) and the milk quality control station (MCS) are also acknowledged for providing data on mastitis incidence.

REFERENCES

Ajzen, I., Fishbein, M. (2005), The influence of attitudes on behavior, In: Albarracin, D., Johnson, B.T., Zanna, M.P. (Eds.) The handbook of attitudes. Lawrence Erlbaum Associates, Inc., Mahwah, New Jersey, 173-221.

Albarracin, D., Johnson, B.T., Zanna, M.P., Kumkale, G.T. (2005), Attitudes:Introduction and scope, In: Albarracin, D., Johnson, B.T., Zanna, M.P. (Eds.) The Handbook of attitudes. Lawrence Erlbaum Associates, Inc. Publishers, MaWah, New Jersey, USA, 3-19.

Andersen, H.J., Enevoldsen, C. (2004), Towards a Better Understanding of the Farmer's Management Practices- the Power of Combining Qualitative and Quantitative Data, In: Andersen, H.J. (Ed.) Radgivning, Bev aegelse mellem data og dialog. Mejeriforeningen, Arhus, 281-301.

Barkema, H.W., Van der Ploeg, J.D., Schukken, Y.H., Lam, T.J.G.M., Benedictus, G., Brand, A. (1999). Management style and its association with bulk milk somatic cell count and incidence rate of clinical mastitis. J. Dairy Sci. 82, 1655-1663.

Barnouin, J., Bord, S., Bazin, S., Chassagne, M. (2005). Dairy management practices associated with incidence rate of clinical mastitis in low somatic cell score herds in France. J. Dairy Sci. 88, 3700-3709.

Barnouin, J., Chassagne, M., Bazin, S., Boichard, D. (2004). Management practices from questionnaire surveys in herds with very low somatic cell score through a national mastitis program in France. J. Dairy Sci. 87, 3989-3999.

Beaudeau, F., vanderPloeg, J.D., Boileau, B., Seegers, H., Noordhuizen, J.P.T.M. (1996). Relationships between culling criteria in dairy herds and farmers' management styles. Prev. Vet. Med. 25, 327-342.

Bergevoet, R.H.M., Ondersteijn, C.J.M., Saatkamp, H.W., Van Woekum, C.M.J., Huirne, R.B.M. (2004). Entrepreneurial behaviour of Dutch dairy farmers under a milk quota system: goals, objectives, attitudes. Agr. Syst. 80, 1-21.

Bigras-Poulin, M., Meek, A.H., Martin, S.W., McMillan, I. (1985). Attitudes, management practices, and herd performance - a study of Ontario dairy farm managers. II. Associations. Prev. Vet. Med. 3, 241.

Chase, L.E., Ely, O.L., Hutjens, M.F. (2006). Major Advances in Extension Education Programs in Dairy Production. J. Dairy Sci. 89, 1147-1154.

Chassagne, M., Barnouin, J., Le Guenic, M. (2005). Expert assessment study of milking and hygiene practices characterizing very low somatic cell score herds in France. J. Dairy Sci. 88, 1909-1916.

Dohoo, I.R., Ducrot, C., Fourichon, C., Donald, A., Hurnik, D. (1997). An overview of techniques for dealing with large numbers of independent variables in epidemiologic studies. Prev. Vet. Med. 29, 221-239.

Dohoo, I.R., Martin, S.W., Meek, A.H. (1984). Disease, production and culling in Holstein-Friesian cows VI. Effects of management on disease rates. Prev. Vet. Med. 3, 15-28.

Elbers, A.R.W., Miltenburg, J.D., De Lange, D., Crauwels, A.P.P., Barkema, H.W., Schukken, Y.H. (1998). Risk factors for clinical mastitis in a random sample of dairy herds from the southern part of The Netherlands. J. Dairy Sci. 81, 420-426.

Field, A. (2005). Discovering Statistics Using SPSS Second Edition, Second edition Edition. Sage Publications Ltd, London, 779 p.

Green, M.J., Bradley, A.J., Medley, G.F., Browne, W.J. (2007). Cow, Farm, and Management Factors During the Dry period that Determine the rate of Clinical Mastitis after Calving. J. Dairy Sci. 90, 3764-3766.

Hutton, C.T., Fox, L.K., Hancock, D.D. (1990). Mastitis Control Practices - Differences between Herds with High and Low Milk Somatic-Cell Counts. J. Dairy Sci. 73, 1135-1143.

Jaccard, J., Blanton, H. (2005), The origins and structure of behavior: conceptualizing behaviour in attitude research, In: Albarracin, D., Johnson, B.T., Zanna, M.P. (Eds.) The handbook of attitudes. Lawrence Erlbaum Associates, Inc., Mahwah, New Jersey, 125-171.

Kaiser, H.F. (1960). The application of electronic computers to factor analysis. Educ. Psychol. Meas. 20, 141-151.

Kiviniemi, M.T., Voss-Humke, A.M., Seifert, A.L. (2007). How do I feel about the behaviour? The interplay of affective associations with behaviors and cognitive beliefs as influences on physical activity behavior. Health Psychol. 26, 152-159.

Kuiper, D., Jansen, J., Renes, R.J., Leeuwis, C., Van der Zwaag, H.G. 2005. Social factors related to mastitis control practices: The role of dairy farmers' knowledge, attitude, values, behaviour and networks. In: Mastitis in dairy production. Current knowledge and future solutions. Proceedings of the 4th IDF International Mastitis Conference, Maastricht, June 2005, 576-582.

LeBlanc, S.J., Lissemore, K.D., Kelton, D.F., Duffield, T.F., Leslie, K.E. (2006). Major advances in disease prevention in dairy cattle. J. Dairy Sci. 89, 1267-1279.

Leeuwis, C. (2004). Communication for Rural Innovation. Rethinking Agricultural Extension. Third Edition. Blackwell Science Ltd, Oxford, 412 p.

Lievaart, J.J., Kremer, W.D.J., Barkema, H.W. (2007). Short communication: Comparison of Bulk Milk, Yield Corrected, and Average Somatic Cell Counts as Parameters to Summarize the Subclinical Mastitis Situation in a Dairy Herd. J. Dairy Sci. 90, 4145-4148.

Likert, R., Hayes, S.P. (1961). Some applications of behavioural research. UNESCO publications center, New York, 333 p.

Nyman, A.K., Ekman, T., Emanuelson, U., Gustafsson, A.H., Holtenius, K., Waller, K.P., Sandgren, C.H. (2007). Risk factors associated with the incidence of veterinary-treated clinical mastitis in Swedish dairy herds with a high milk yield and a low prevalence of subclinical mastitis. Prev. Vet. Med. 78, 142-160.

Reneau, J.K. 2001. Somatic cell counts: measures of farm management and milk quality. In: National Mastitis Council Annual Meeting Proceedings (2001), National Mastitis Council, Verona, Wisconsin, 29-37.

Reneau, J.K. 2002. Milk Quality Mind Set. In: Proceedings of the Great Lakes Dairy Conference, Oregon, Ohio, April 30- May 2 2002.

Seabrook, M.F. (1984). The Psychological Interaction between the Stockman and His Animals and Its Influence on Performance of Pigs and Dairy-Cows. Vet. Rec. 115, 84-87.

Tarabla, H., Dodd, K. (1990). Associations between farmers' personal characteristics, management practices and farm performance. Brit. Vet. J. 146, 157-164.

Vaarst, M., Paarup-Laursen, B., Houe, H., Fossing, C., Andersen, H.J. (2002). Farmers' choice of medical treatment of mastitis in Danish dairy herds based on qualitative research interviews. J. Dairy Sci. 85, 992-1001.

Van der Zwaag, H.G., Schaik, G., Lam, T.J.G.M. 2005. Mastitis control programs in The Netherlands: Goal, tools and conditions. In: Mastitis in dairy production. Current knowledge and future solutions. Proceedings of the 4th IDF International Mastitis Conference, Maastricht, june 2005, 599-600.

Wenz, J.R., Jensen, S.M., Lombard, J.E., Wagner, B.A., Dinsmore, R.P. (2007). Herd management practices and their association with bulk milk somatic cell count on United States dairy operations. J. Dairy Sci. 90, 3652-3659.

Willock, J., Deary, I.J., McGregor, M.M., Sutherland, A., Edwards-Jones, G., Morgan, O., Dent, B., Grieve, R., Gibson, G., Austin, E. (1999). Farmers' attitudes, objectives, behaviors, and personality traits: The Edinburgh Study of Decision Making on Farms. Journal of Vocational Behavior 54, 5-36.

COMPARATIVE TRANSMISSION DYNAMICS OF 6 ENDEMIC INFECTIONS IN 114 CATTLE HERDS OVER 3 YEARS: PRELIMINARY OBSERVATIONS

MEDLEY, G.F., WOODBINE, K.A., RAMIREZ-VILLAESCUSA, A.M., MOORE, S.J. AND GREEN, L.E.

SUMMARY

The farm-level dynamics of six endemic infections of cattle in the UK are compared. Longitudinal data over three annual visits are presented showing the prevalence of antibodies to bovine viral diarrhoea virus (BVDV), *Mycobacterium avium* spp. *paratuberculosis* (MAP), *Neospora caninum* (NC), *Leptospira hardjo* (LH) and bovine herpesvirus type-1 (BHV). Additionally, the numbers of reactors to bovine tuberculosis (BTB) are presented. Farms differed appreciably in the variability of serology between visits. Of the six infections, BVDV appears the most likely to produce epidemics, and the most likely to persist by "jumping" from herd to herd. Seroprevalence of NC and MAP is remarkably constant, especially given that 50% of the cattle sampled in the first visit had been replaced by the third visit. The herd-level relationships between the six infections are presented. The potential mechanisms important in determining the persistence of these infections is discussed.

INTRODUCTION

Endemic infectious diseases are the major cause of production loss and welfare concern in cattle, but are relatively low priority in terms of control and research. In contrast, epidemics of exotic pathogens (e.g. foot and mouth disease virus in the EU) attract more attention, are more vigorously controlled (eliminated), and consequently relatively much better researched and therefore understood. By definition, endemic diseases persist in the host population, and the central aim of this paper is to describe the dynamic patterns for six different infections persisting within the UK cattle population.

Generally, there are a number of distinctions that can be drawn between pathogens in terms of the processes that enable their persistence. For example, some pathogens are able to persist in the environment (used broadly to include wildlife reservoirs), and some can only survive in cattle. Pathogens which use the environment can be thought of as persisting on farms (rather than herds). The duration of infectiousness in an individual bovine host is another important characteristic that influences persistence. Infections that are short-lived must have access to a ready supply of susceptible animals in order to persist; infections that induce, for example, a carrier state are able to "wait" for susceptible animals. Consequently, there is the possibility of developing a generic framework of endemic disease and its mechanisms of persistence (Green, 2007).

From a theoretical stance, farms can be thought of as metapopulations, i.e. a population of populations that are connected through animal movement. There is a continuum of possibilities

for persistence within a metapopulation. First, an infection may persist long-term within individual populations (farms), either in the environment and/or in the cattle. Transmission occurs when susceptible cattle arrive on infected farms through movement (purchase) or birth. If the infection persists for a long period within individual cattle, then movement of infected cattle between farms will result in transfer of infection between farms that then might spread within the herd if there is onward transmission.

Second, an infection may only persist within the herd (i.e. there is no environmental reservoir) and the duration of infection within individual cattle may be relatively short. Introduction of infection into a susceptible herd will result in an epidemic and loss of infection from the herd either through natural "fade out" or by movement (sale) of infected cattle. After a period during which susceptibility in the herd increases, the herd is then susceptible to infection until the infection is re-introduced, and another epidemic ensues. In this extreme, the infection persists by movement between susceptible herds.

In reality, of course, most infections will sit between these two extremes, but might be predisposed towards one pattern. Figure 1 illustrates these two different processes. The first observation is that these patterns can only be distinguished by longitudinal studies. The results from a cross-sectional survey (represented by the vertical dotted lines in Fig. 1) will not show the different dynamic patterns. The second observation is that whilst studies to determine risks for farm-level infection are valid if farms are constantly infected (left panel), they are less useful if the farms that are infected are changing with time (right panel). In this second situation, the risks for infection are dependent on the current situation (e.g. recent introduction of infection), the history of the farm (e.g. the farm must have been free of infection for sufficient time for susceptibility to develop so that the introduction is successful) and the infection status of other (source) farms. Such "misclassification" will result in reduced statistical power to detect real risks.

It is also worth considering the timescales over which farms will vary their infection status. If there is no environment component, then the turnover of infection within the herd will be dominated by the shorter of the infection period or duration that each animal stays on the farm. Thus, farms are expected to change state more frequently for a pathogen with a short infectious period than a pathogen with a long infection period or one with an environmental reservoir. Carrique-Mas *et al*. (2008) have separated the effects of farm and herd in terms of risks of herd breakdown with bovine tuberculosis. The foot and mouth epidemic (FMD) in 2001 resulted in complete removal of cattle from some farms, but the effect of infection on the farm before FMD remained after FMD. There was also a demonstrable risk related to purchase of cattle, demonstrating that both environmental contamination and purchase of infected cattle are important in determining herd status.

Even in the situation when the infection status of farms is roughly constant, it should be remembered that the farm states are not independent (i.e. animals are moving between farms), and the composition of the herds is continually changing as animals move in and out. Movement of infected animals on to infected farms might be important for pathogen population genetics (e.g. the movement of different antigenic types or genetic resistance). It might also prevent infected farms losing their infection through natural fade out. Within herd prevalence of infection will also be affected by movement, being both increased and decreased. Within-herd transmission might be determined by factors that are different from those determining the infection status of the farm, and it is quite possible to imagine processes (such as purchasing and selling) that would increase the risk of farm infection but reduce the prevalence on infected

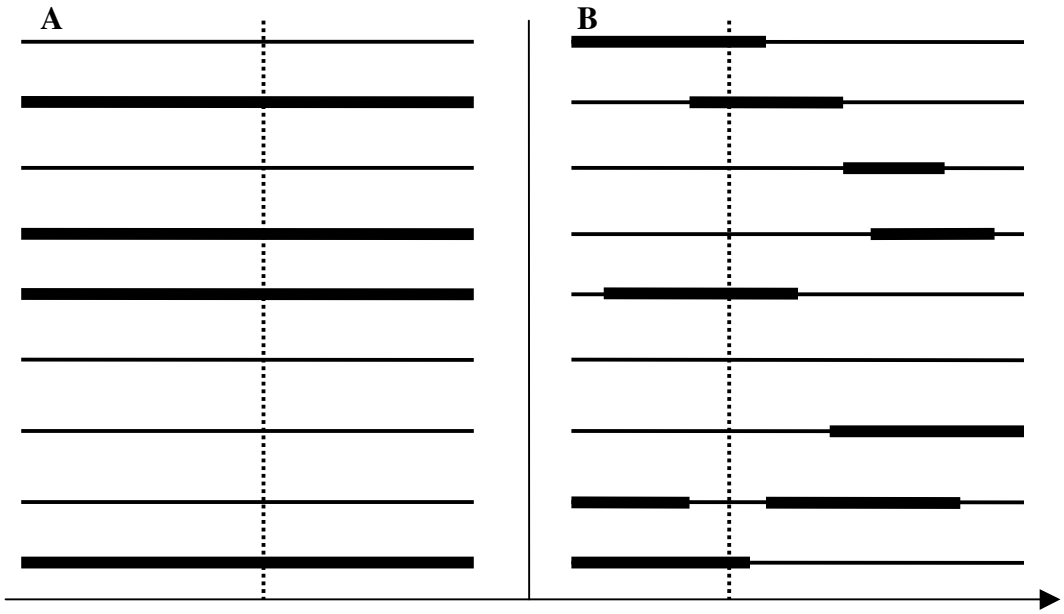


Fig. 1. Schematic diagram of persistence of infection in a metapopulation over time. Left panel (A) shows the situation with 4/9 herds infected constantly. Right panel (B) shows intermittent infection with similar herd level prevalence but different temporal dynamics. The vertical dotted lines represent cross-sectional studies conducted at one point in time.

farms. Between farm and within herd prevalence are different, but related processes, and require careful consideration.

Results are presented from a cohort study of 114 farms in south west England. The study protocol was developed on the basis of understanding the cattle-cattle transmission of bovine tuberculosis (BTB) (see Ramirez-Villaescusa et al., 2008, this volume). On farm visits, we also took a sample of blood (stored as serum) from all accessible adult cattle on the farms, and have used this to derive data relating to five further infections: bovine viral diarrhoea virus (BVDV), Mycobacterium avium spp. paratuberculosis (MAP), Neospora caninum (NC), Leptospira hardjo (LH) and bovine herpesvirus type-1 (BHV). This choice was on the basis of several considerations. First, they cross several spectra in terms of phylogeny, and natural history. Some have a known environmental reservoir (BTB, MAP, NC, LH) compared with none (BVDV, BHV); to compare relatively short infectious period (LH), chronic infection (BTB, MAP, NC), recrudescence (BHV), a defined carrier state (BVDV – persistently infected cattle); and to compare infections with known vertical (maternal) transmission (MAP, NC, BVDV) with infections where vertical transmission is not suspected (BTB, BHV, LH).

Second, serology is available and the natural history and pathology of the infections is understood – albeit to different resolutions. We have access to the BTB testing history. Third, taken together, with BTB, these infections represent the major production constraints imposed by endemic disease in the UK (Bennett et al 1999; Chi et al, 2002) – MAP, BHV, BVDV and BTB are all potential restrictions to international trade.

Detailed analyses of the data for each infection will be presented in separate papers currently in preparation. The principal aim of the current paper is to compare the farm-level dynamic patterns.

MATERIALS AND METHODS

The data used in this study came from a 4 year cohort study of 114 cattle (dairy and grower/suckler herds) in south west England that took place from 2002 – 2006. All farms were situated in areas within the Randomised Badger Culling Trial (RBCT) that was conducted in England from 1998–2005 (Bourne et al., 2007), and in an area where some herds were restocked (i.e. completely depopulated and subsequently restocked) after the 2001 foot-and-mouth disease (FMD) epidemic.

Each farm was scheduled to be visited three times, approximately annually, and a blood sample taken from all accessible adult (>2yrs old) cattle that were expected to be present at a subsequent visit (i.e. breeding stock). Samples (up to 10ml) were collected under Home Office licence. A further subset of herds (n = 15) were re-visited and blood samples were taken from cattle of all ages. These herds were re-visited either to re-test individual cattle to confirm whether they were persistently infected (PI) with BVDV or to sample the whole herd (including young stock) after a BVDV-PI had been detected in the adult herd. Four herds had a whole herd test instead of a routine third visit because a BVDV antigen positive sample was detected before this third visit. Blood samples were centrifuged at the University of Warwick at 3220g for 15 minutes and serum was removed, frozen and stored at -20°C until they were analysed. Farmers were interviewed between June 2003 and February 2004 using a standardised questionnaire to obtain information about clinical disease. Participation in the questionnaire was >95%.

In total, 29,782 blood samples were tested from 15,736 cattle in 114 herds. There were 9963, 8979 and 8580 samples respectively from the 114 herds visited once, 102 visited twice and 96 visited three times. An additional 1135 samples were collected when four whole herd visits replaced the routine third visit (i.e. extra samples taken from young stock <2 years). A further 1125 samples were taken at additional follow-up visits.

Cattle ear tag or freeze brand numbers were recorded during each visit; when a freeze brand was taken the farmer provided a list describing the associations between freeze brand and ear tag numbers. Upon receiving the samples at the University of Warwick each ear tag was matched with information from the Cattle Tracing System (CTS) describing the cattle date of birth, origin (whether it was homebred or purchased), breed and sex. One percent (371 samples) of the ear tags did not match with the CTS data. Ninety percent of these cases were because the same identifier had been recorded twice; other errors were that the cattle did not have a freeze brand or ear tag. Information provided by the British Cattle Movement Service (BCMS) allowed us to track cattle movements over time and match dams and calves. The information from the BCMS was matched using the ear tag numbers and blood sampling dates; for cattle born after 2001, at which point it became compulsory to record all cattle births, >99% of cattle were in the BCMS dataset.

The sera were tested for presence of antibodies against all infections apart from BTB. Sera that were negative for BVDV-Ab were tested for bovine viral diarrhoea virus antigen (BVDV-Ag). All serology was conducted with purchased kits and carried out according to manufacturers' instructions. All samples were tested in duplicate and retested if there was a significant discrepancy between duplicates. Detailed information was also collected on the incidence of reactors to routine bovine tuberculosis skin tests.

Laboratory results, questionnaire data and external data were entered into a relational database (PostgreSQL, PostgreSQL Global Development Group) using Microsoft Access

(Microsoft Corp. US) as a front end. All data were checked for errors and data re-entered where errors were detected.

Analyses of these data are currently being done, and will be reported. Here, the data are used to derive an understanding of the processes that might be important considerations in considering the mechanisms by which these endemic infections persist in the UK cattle population.

RESULTS

Time-Dependent Patterns

The antibody results have a different interpretation for each infection. For NC, BHV and MAP, they indicate presence of infection, since cattle commonly develop a chronic infection state. For BVDV and LH, the presence of antibodies does not necessarily indicate presence of infection, but the changes over time do: the increase in BVDV prevalence from one time period to the next is indicative of transmission (i.e. new infection) during that period.

LH appears to be unusual, in that none of the variables that we have investigated (i.e. farm covariates), apart from age, were significantly related to individual or farm-level prevalence. It would appear to be a genuine environmental pathogen, with roughly equal risk of infection for all cattle. Consequently, throughout the paper we use LH as the null hypothesis, i.e. the patterns that would be expected if there is no cattle-cattle transmission and the environmental contamination is independent of cattle infection.

Figure 2 shows the within-herd prevalence of antibodies to four infections over three consecutive visits (approximately 3 years). The pattern for MAP is the same as that for NC, and is not shown. Farm-level prevalence does not vary greatly over the period of the data for NC (and MAP). The small number of herds in which prevalence varies >20% demonstrates that such dramatic changes can occur through infection and cattle movement. Although these farms are a cohort, it should be remembered that the cattle population within the farms changes fairly rapidly. Of the 9,887 cattle sampled in the first visit, only 6,490 (66%) and 4,387 (44%) remained on the farm for the second and third visits respectively, i.e. the median residence for adult cattle is between 1 and 2 years. Consequently, a stable herd prevalence, as seen for NC and MAP, does suggest a strong farm effect.

In contrast the patterns for BHV and BVDV are more variable. They are similar to each other, although prevalence of antibodies to BVDV is generally much higher. The epidemics of BHV are consistent with intermittent shedding by seropositive cattle. But, note the herds in which BHV remains constant, at both low prevalence (only two herds with <20% prevalence at the first visit increased above 20% at the subsequent visits) and high prevalence (only four farms with >60% prevalence at the first visit decreased below 60% at subsequent visits). These two observations suggest that the transmission of BHV is highly heterogeneous.

Dramatic increases in BVDV prevalence are more likely observed in herds with low BVDV seroprevalence, and epidemics of BHV in herds with intermediate BHV seroprevalence. This is consistent with BVDV epidemics resulting from introduction of BVDV into a susceptible adult herd. There were 40 antigen positive cattle in 26 herds, of which 11 were confirmed persistently infected cattle, 16 were probable PI cattle and 13 probable transiently infected cattle

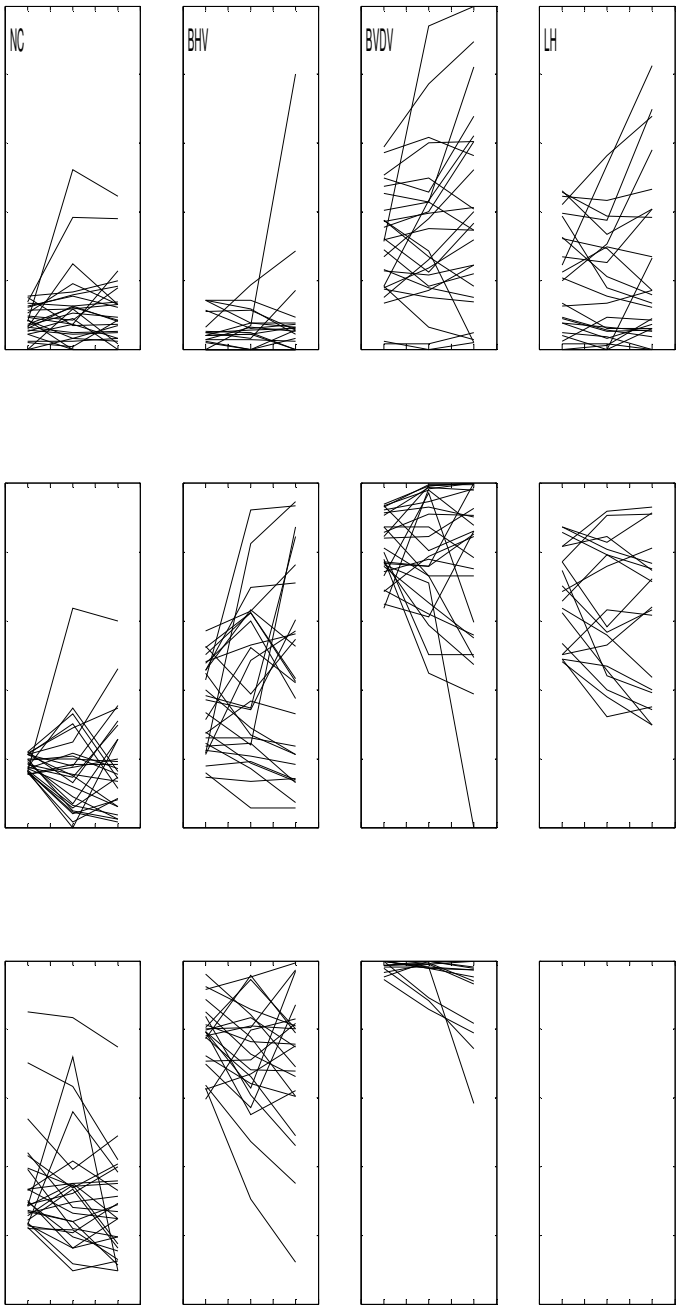


Fig. 2. The panels show the prevalence of infection over three visits (horizontal axis) – each farm is represented by one line. The vertical scale is 0-50 for NC and 0-100 for BHV, BVDV and LH. The farms have been divided into three equal group sizes with ascending prevalence on first visit (top to bottom rows). Farms with less than 20 cattle at any visit are not shown. Farms are not shown for the infection against which they are vaccinating. One third of farms were vaccinating against LH resulting in only two groups.

Effect of Herd Size

Figure 3 shows the mean prevalence of infection as it varies with mean herd size. This is the crude, unadjusted relationship. Multivariate analyses show that BVDV, LH and NC have no relationship with herd size, whereas MAP, BHV and BTB all have increased prevalence as the number of cattle on the farm increases.

Herd size is a proxy measure for many aspects of on-farm cattle demography. Cattle density might be increased by herd size, but not necessarily. Birth rates, cattle movement and farm type (dairy, beef etc.) are both closely correlated with herd size, and both impact directly on the supply of susceptible cattle and the introduction of infection into the herd. Of the 15,736 cattle sampled, 6,751 (43%) were purchased (i.e. we sampled them on a herd different from their natal herd) and 7,679 (49%) were home-bred cattle; herd of origin information was missing for 1306 (8%) cattle. Thus, at any one time, approximately 50% of the cattle population has moved between at least two farms at some point previously.

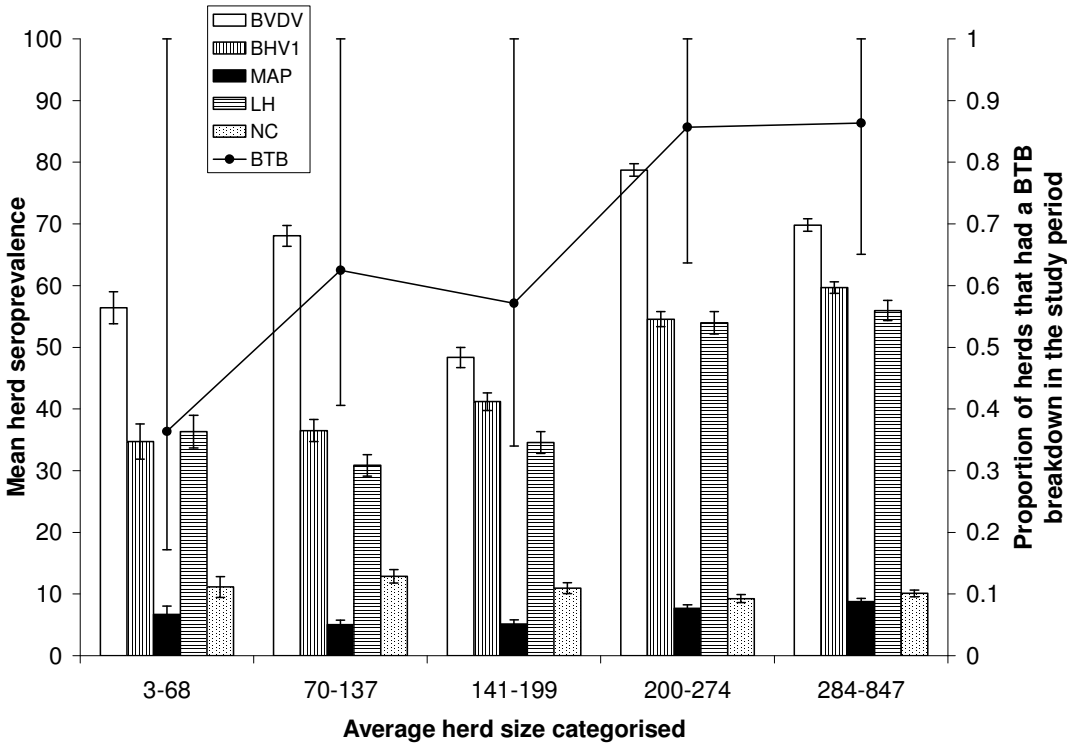


Fig. 3. The relationship between herd size and the prevalence of six endemic infections. The seroprevalence is on the left axis, the proportion of herds with BTB reactors on the right.

Relationship between Infections

Table 1 shows the Spearman rank correlation coefficients between seroprevalence of five infections at the three visits. Purchased cattle have been excluded to emphasise the farm state. If purchased cattle are included then the only significance is a positive correlation between BVDV and LH. There is a consistent positive correlation between prevalence of antibodies to LH and BVDV. Many herds (38/114, 33%) were vaccinating against LH, so much of the high seroprevalence for LH will not be related to infection. However, this would suggest that

vaccines against LH are being used in herds with higher prevalence of BVDV in cattle born on that farm, which is indicative of recent or current transmission of BVDV on the farm.

The tables also demonstrate that the correlations between infections change over time – a single cross-sectional survey would have been misleading. There appears to be a consistent LH-BVDV-BHV relationship. Interestingly, the relationship between NC and MAP only appears in the third visit. Given the apparent farm level effect (i.e. independent of the herd) seen in the consistency of the seroprevalence over time (Fig. 2), it might have been expected that the NC-MAP correlations would have been consistent. This result deserves further exploration.

Table 1. Correlations between herd seroprevalence of infections assessed at three different visits in home-bred cattle only

	BVDV	**BHV**	**MAP**	**NC**	**LH**
Visit 1					
BVDV	**—**				
BHV	0.0277	**—**			
MAP	0.0531	0.0992	**—**		
NC	-0.0841	-0.0315	0.0035	—	
LH	**0.2099 (0.0263)**	**0.1706 (0.0722)**	0.0122	-0.0894	—
Visit 2					
BVDV	—				
BHV	**0.1952 (0.0541)**	—			
MAP	0.0597	0.0963	—		
NC	-0.0341	-0.1018	0.1443	—	
LH	**0.1913 (0.0592)**	**0.2079 (0.0400)**	0.0768	-0.1154	—
Visit 3					
BVDV	**—**				
BHV	**0.2107 (0.0404)**	**—**			
MAP	0.0701	**0.2595 (0.0111)**	**—**		
NC	-0.0658	0.0530	**0.2349 (0.0219)**	**—**	
LH	**0.3274 (0.0012)**	**0.1819 (0.0778)**	0.0882	-0.0441	**—**

Numbers (p-value) in bold are significantly or marginally significant.

Figure 4 shows the mean herd seroprevalence for herds that did and did not have any BTB disclosed on the farm by routine surveillance during the period of the study. There were positive relationships between BTB breakdowns and seroprevalence of BHV and LH. Again, this might be due to vaccination in herds more likely to have BTB.

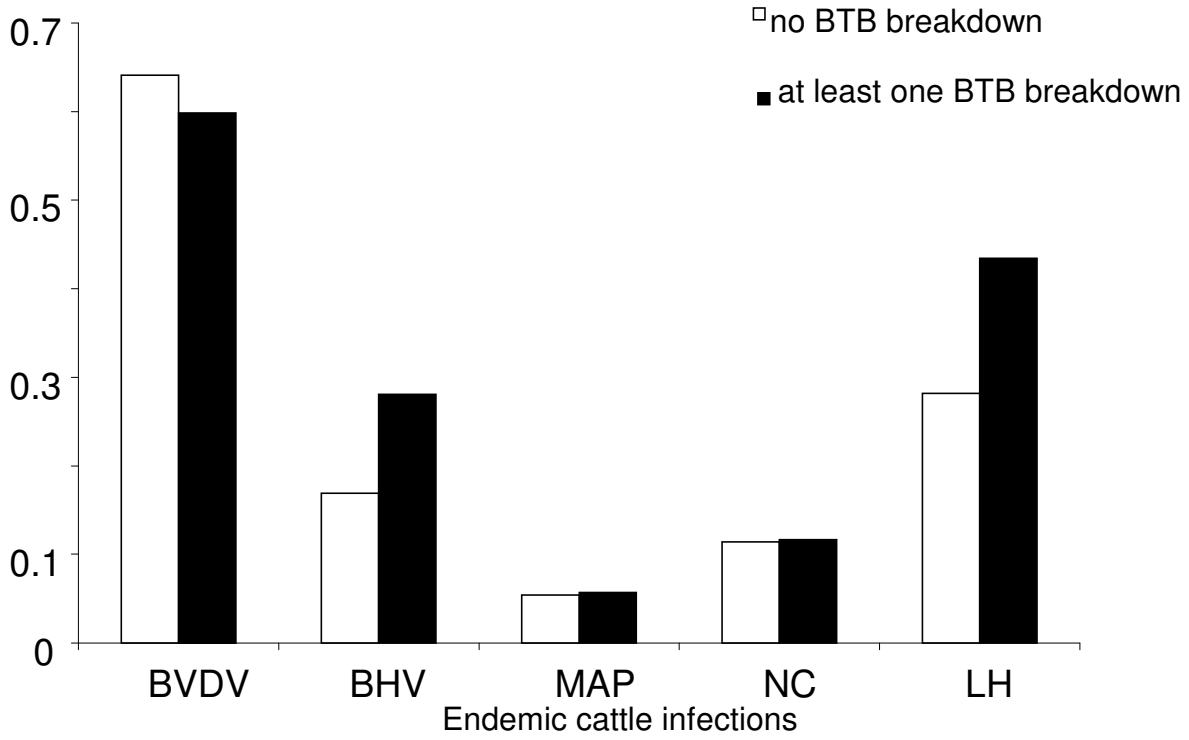


Fig. 4. The relationship between the mean seroprevalence of five pathogens and risk of having a BTB reactor in the herd during the study. The relationships with LH ($p = 0.0385$) and BHV ($p = 0.0272$) are significant by Wilcoxon rank sum.

DISCUSSION

Endemic infections of cattle persist in the whole population by transmitting both within herd and between herds, and potentially infecting premises (farms). Longitudinal data are required to understand these mechanisms. Data are presented illustrating patterns over a period of three years for the six infections.

The constant farm-level prevalence for NC and MAP is expected for pathogens with significant environmental reservoirs. However, given the degree of cattle movement between farms, this consistency is remarkable; it requires that antibodies to NC and MAP are markers of recent exposure, i.e. they decay with time since exposure, and consequently the prevalence on a farm indicates the degree of exposure on the farm. Only six farms had a NC seroprevalence >20% on at least two visits, suggesting that this infection (and MAP) persist through infecting farms rather than herds. Interestingly, these infections both have recognised vertical transmission potential; although this should not particularly result in constant farm prevalence (i.e. purchase of cows infected with NC or MAP should result in infected calves), it might interact with buying and selling management to induce this pattern.

The seroprevalence of LH, BVDV and BHV vary on the timescale of duration of residency of cattle on the farm, since antibodies are not lost during the lifetime of an individual. The change in observed seroprevalence over time is determined by both the seroprevalence itself (i.e. a lower seroprevalence indicates susceptibility) and the presence of the infection.

The interpretation of the patterns differs for these infections. The epidemiology of LH appears to be independent of cattle infection, and that cattle are continually at risk of infection. The presence of antibodies to BVDV does not necessarily imply the presence of infection. The presence of antibodies to BHV indicates the presence of infection, since, as a herpes virus, infection is rarely cleared from an infected individual. Thus, herds with intermediate BHV seroprevalence are more likely to show an increase: they have both the infection and the susceptible cattle to infect. In contrast, epidemics of BVDV are more likely in herds with low BVDV seroprevalence, when infection is introduced into the adult herd. Over the period of the data six clear epidemics of BVDV were observed, and BVDV appears to persist in the cattle population by jumping from herd to herd more than the other five infections. We intend to use model-based approaches to estimate the parameters of transmission (e.g. Viet & Medley, 2006).

The relationships with herd size might be related to the density of cattle for an environmental pathogen; however, it could also be related to purchasing and selling behaviour. The contrasts between BVDV and BHV and between NC and MAP are interesting since these have similar dynamic patterns, but potentially different mechanisms determining these patterns since they have different relationships to herd size.

Of the six infections studied, BVDV appears to most likely to produce epidemics, and the most likely to persist by "jumping" from herd to herd. Over the three visits, six epidemics of infection were clearly seen from the serology.

The relationships between the farm-level prevalence of different infections demonstrate, again, the importance of considering the patterns as dynamic. The correlations including all farms show very little correlation, apart from between BVDV and LH, which remains in all sub-divisions of the data. However, these results are clouded by vaccination – the vaccination status of purchased cattle is unknown; farm-level information was recorded, but the status of individual home-bred cattle was not recorded. Future analyses will include the relationship at the level of the individual animal. Currently, the most pragmatic conclusion from the farm-level data is that these infections are essentially independent.

Management variables, for example the types of herd (dairy, grower/suckler etc), are critical for designing intervention programmes to reduce the burden of endemic disease. Undoubtedly, these will explain some of the patterns and relationships observed. For example, the LH-BVDV-BHV relationship is probably related to the increased prevalence of antibodies to these infections in dairy herds. However, we believe that, in initially at least, it is important to focus on those variables that are most amenable to farmer control, and most likely to have a causal mechanism in relation to endemic disease (e.g. cattle age and movement) in order to explain the epidemiology of endemic infection.

ACKNOWLEDGEMENTS

The fieldwork and sample collection was funded by DEFRA (SE3026). The laboratory testing and data analysis were funded by BBSRC (BBS/B/04854). The fieldwork and laboratory testing was conducted by a team of technical staff, whose professional contribution is gratefully acknowledged. We thank the farmers for access to their cattle and data.

REFERENCES

Bennett, R.M., Christiansen, K. and Clifton-Hadley, R.S. (1999) Preliminary estimates of the direct costs associated with endemic diseases of livestock in Great Britain. Prev. Vet. Med. 39 (3), 155-171.

Bourne, J. et al., (2007) Bovine TB: The Scientific Evidence. Final Report from the Independent Scientific Group, Defra (http://www.defra.gov.uk/animalh/tb/isg/index.htm).

Carrique-Mas, J.J., Medley, G.F. and Green, L.E. (2008) Risks for bovine tuberculosis in British cattle farms restocked after the foot and mouth disease epidemic of 2001. Prev. Vet. Med. (in the press).

Chi, J., VanLeeuwen, J.A., Weersink, A., Keefe, G.P. (2002) Direct production losses and treatment costs from bovine diarrhoea virus, bovine leucosis virus, Mycobacterium avium subspecies paratuberculosis, and Neospora caninum. Preventive Veterinary Medicine 55, 137-153.

Green, L.E. (2007) Improving farm animal health – understanding infectious endemic disease. Society for Veterinary Epidemiology and Preventive Medicine Proceedings 2007, pp13-25.

Ramirez-Villaescusa, A., Medley, G.F., Mason, S. and Green, L.E. (2008) Risk factors bovine tuberculosis skin test reactivity for 48,055 cattle on 148 herds in south west England. Society for Veterinary Epidemiology and Preventive Medicine Proceedings 2008.

Viet, A.F. and Medley, G.F. (2006) Stochastic dynamics of immunity in small populations: a general framework. Mathematical Biosciences 200, 28-43.

IMPACT OF PRIOR EXPOSURE TO BOVINE TUBERCULOSIS ON THE RISK OF BOVINE TUBERCULOSIS SKIN TEST REACTIVITY FOR 48,055 CATTLE IN 148 HERDS IN SOUTH WEST ENGLAND

A. RAMIREZ-VILLAESCUSA, G.F. MEDLEY, S. MASON AND L.E.GREEN

SUMMARY

The aim of this study was to investigate the effect of movement factors on the risk of a bovine animal becoming a reactor at a bovine tuberculosis (bTB) herd skin test. Using data from 148 study farms, individual animal records were created by interrogating the national movement database and the bTB testing databases, and were corroborated with data collected at farm visits. Study farms, 24% of which were restocked due to the foot and mouth disease (FMD) outbreak in 2001, were located in six counties of the south west of England. All farms participated in the Randomised Badger Culling Trial (RBCT) trial areas (the Kreb's trial). Cattle were more likely to react to the bTB skin test when they had been present at a previous test or tests where a bovine / cattle had reacted to the skin test, highlighting the importance of infected cattle as a main source of infection to other cattle on the same farm or another farm if moved when infected but undetected.

INTRODUCTION

Bovine tuberculosis (bTB) is an infectious disease of cattle caused by *Mycobacterium bovis* (*M. bovis*). In the south west of England, the pathogen has been endemic for many years.

There is evidence, from experimental studies, that cattle to cattle transmission does occur (Pollock et al. 2006). This may occur on farms in the UK because infected cattle are not detected either because they are not tested or because truly infected cattle test negative because of the relatively low sensitivity of the skin test. The test is estimated to be more than 99% specific, with a sensitivity of approximately 74% to 95% (Monaghan et al. 1994; Costello et al. 1997). Cattle movements play an important role in determining the frequency of testing individual animals because movement can result in cattle missing tests or being tested more frequently. Testing is regular (at 1,2 or 4 year intervals) within regions of GB, but not random, with up to 80% of cattle in GB not tested in their lifetime (Mitchell et al. 2006). Cattle herds in the south west of England (and all herds in the current analysis) are tested for bTB annually with the comparative cervical tuberculin test (SICCT).

The current study was conducted to investigate the impact of prior exposure to bovine tuberculosis (defined as an animal present in a herd and of the correct age to be tested when a reactor was detected). Movement of cattle was included in the analysis and cattle might have been exposed in the study herd or in a previous herd. This is a proxy for potential cattle-cattle transmission. Earlier field epidemiological studies (mostly case-control studies) considered herd

as the outcome of interest. In the current study, the outcome was whether or not a bovine became a reactor at a herd test.

MATERIALS AND METHODS

General methods

The reference farms

The reference farms were located in areas where the Randomised Badger Culling Trial (RBCT) was undertaken (Kreb's et al. 1997; Bourne et al. 1999). The areas were in the South West of England (Cornwall, Devon, Somerset and Gloucestershire) and the West Midlands (Herefordshire and Worcestershire). Within the RBCT, there were thirty trial areas, of approximately 100 km² each. The areas were identified as ten triplets (A to J) or geographical areas, with three treatments (reactive culling of badgers, proactive culling of badgers or surveillance of badgers only) within each triplet. These farms were also located in areas where the foot-and-mouth disease (FMD) epidemic of 2001 had led to depopulation of cattle farms (Gibbens et al. 2001).

The study period

The study period for herd breakdown with bTB was from 1st June 2001 to 19th August 2004. The exposure of cattle to bTB was from birth, in cattle with dates of birth recorded in the BCMS. This was from 1996 (see below).

The study farms

Study farms were recruited between November 2002 and October 2003.. Farms that were restocked after being depopulated because of FMD in 2001 were defined as exposed, and those that were continuously stocked were defined as non-exposed. Exposed farms were selected from a population sample of all cattle farms that were depopulated because of FMD whilst non-exposed farms were randomly selected from all cattle farms within the trial areas: one restocked farm was matched with three continuously stocked farms within an RBCT trial area. There was a distance of at least 1 km between the restocked and each of the continuously stocked farms to ensure that cattle-cattle transmission was not confounded by nose-nose contact over farm boundaries.

Each trial area and study farm was identified by codes, but the nature of badger control was not disclosed until the study was completed.

Study animals

All cattle present on the study farms at each of the herd tests conducted between 1st June 2001 and 19th August 2004 (see details below, including exclusion) were enrolled into the study.

Data collection and management

Data collection

Data were extracted from two sources: the Cattle Tracing System (CTS) of the British Cattle Movement Service (BCMS) and the VetNet databases. Records from the BCMS were available from the 1st July 1996 to the 4th August 2004 and from VetNet from 1st January 1995 to 19th August 2004. Reactor animals were identified by matching reactor data from the VetNet database with animal identity data from the BCMS database. It was not possible to achieve exact matches with approximately 60% of reactors, however, in almost all situations this was due to minor recording errors which were rectified.

a. The CTS database: The Cattle tracing System database of the BCMS records all the movements of cattle registered or imported into GB. These include movements on and off farm, to and from markets, and to abattoirs. Farms are identified by County Parish Holding (CPH) and cattle by unique ear-tag numbers. The information used from the database included: date of birth, sex, breed, date of movement (if purchased on to the study farm), date of movement off the study farm (if this occurred) and all movements, if any, before a bovine joined the study herd. Cattle were defined as either purchased or born on the study farm, and had either left or were still on the study farm at the end of the study period. For cattle that had left the farm before the end of the study, the length of time spent on the farm was defined by the date the animal was born or moved onto the farm to the date that it had died or was moved off the farm. Cattle were assumed to be on the farm at the end of the study if there were no off-farm movement records, and were assumed to have been tested if they were present on the farm when a herd test took place, with the exception of those cattle excluded because of the type of test e.g. too young to be tested. Before any tests were excluded, a total of 161,782 cattle tests were identified in the database of study cattle. Where errors in the data were obvious, these were corrected.

b. The VetNet database: The VetNet database provides records of the results of the SICCT test. Herd and animal test databases were used. There are two types of test: those that target whole herds (herd tests) and those that target individual cattle (animal tests). Using the current test type definition and test criteria (DEFRA 2005), individual animal tests such as VE-IR, VE-PII, VE-PRI, VE-SLH or VE-TR, and herd tests such as VE-6M, VE-12M, VE-WHT or VE-WHT2 when bovine animals were calves under six weeks old, were excluded. VetNet database does not have records for cattle that tested negative to the SICCT test, hence the need to use information from the CTS database to establish cattle that were likely to have been tested.

Data management

The data were managed using Access, Microsoft Corp. US. A dataset was created with a hierarchical structure arranged in three levels: tests (level 1), animals (level 2) and herds (level 3).

Statistical analysis

All farms with at least one herd test were used in the analysis during the study period (consequently four farms were excluded).

The outcome variable was binary: a bovine animal was a reactor or not at a test. Although an animal could be tested several times, it could be a reactor only once, due to its removal from the farm if disclosed as reactor. Logistic regression with random effects was carried out using the statistical package MLwiN, Version 2.0 (Rasbash et al. 2004). The analysis was

implemented by using a Generalised Linear Mixed Model (GLMM) with a 3-level hierarchical structure. In MLwiN, 1st order marginal quasi likelihood (MQL) estimates were derived using Iterative Generalised Least Square (IGLS).

The model took the form

$$p_{ijk} \sim \beta_0 + \sum \beta X_k + \sum \beta X_{kj} + vk + u_{kj} + e_{ijk} \quad (1)$$

Where p_{ikj} = a bovine reactor yes / no, at test ijk, for animal jk from farm k ~ investigated with a logit link function, β_0 = intercept, and βX is series of vectors of fixed effects varying at level 1 (ijk) or level 2 (jk) or level 3 (k), and $v_k + u_{kj} + e_{ijk}$ are the level 3, 2 and 1 residual variances.

Twenty seven explanatory variables were screened using univariate analysis. The variables were test variables (e.g. animal age) animal variables (i.e. sex, breed, born or purchased on study farm) and herd variables (i.e. RBCT treatment, restocking status, bTB history of the herd).

The potential exposure to other reactors was defined using a categorical variable that combined whether cattle were born on the study farm or purchased and whether cattle had never been in a test with a reactor in the past (yes, no or untested) on the study farm and for purchased cattle for tests on up to two previous two farms.

After initial screening, the eleven trial areas from the RBCT represented in the study were grouped into the three treatments (reactive, proactive, survey only). The type of test was divided into three groups: yearly, if it was a routine test; short-interval (VE-SI) if carried out sixty days after a previous test and other strategic tests if tests were carried out with a control purpose and more frequently than yearly, for example, a check test due to a slaughterhouse case. Yearly tests included tests coded as VE-WHT, VE-RHT, VE-12M and VE-CON12 and strategic tests included tests coded as VE-CT, VE-6M, VE-CON and VE-CON6.

Once the final model was built Monte Carlo Markov Chain (MCMC) with Gibb's sampling was used to reduce the conservative estimates of the standard errors. The model was run for 70,000 iterations and a burn-in period of 5,000.

The precision of the parameter estimates was assessed by using kernel density plots. The model was run with a data point with a large residual value (a herd with 78 reactors at a test) absorbed in the model as dummy variable and checked that the estimates did not change their value significantly. The observed and expected values were divided into deciles. Model fit was then assessed by calculating the Pearson's Chi-square (Hosmer and Lemeshow 2000).

RESULTS

Descriptive results

The study farms

A total of 148 study farms, in six counties of the South West region, were recruited into the study. The farms were located in six (A, B, C, H, I and J) of the ten RBCT triplets, and eleven of the thirty treatment areas: 37% were in the reactive treatment, 28% in proactive and 35% in survey only.

During the study period, 697 herd tests were conducted on 144 study herds (including 110 continuously stocked and 34 restocked farms); the first test was carried out on 25th June 2001 and the last on 3rd August 2004. Four of the 148 study herds were not tested with a herd test during this period; two of these had animal tests and the other two herds were last tested before June 2001. The number of herd tests per year varied: 4% of the 697 tests were carried out in 2001, 29% in 2002, 45% in 2003 and 22% in 2004. The decrease in the number of tests in 2001 was due to the disruption caused by the FMD outbreak in 2001. Consequently, most tests due in 2001 were not completed. In total, 564 (80.9%) and 133 (19.1%) herd tests were conducted in continuously stocked and restocked herds, respectively. During the study period, the median number of tests on continuously stocked and restocked herds was 4 and 3, respectively.

The study animals

At the 697 herd tests in the 144 study herds, 156,562 animal tests were conducted on 48,055 cattle, and 723 reactors were disclosed. A further 53 reactors were disclosed in these herds during this period, but not at herd tests. The percentage of cattle tested, by restocking status and RBCT trial area, is presented in Figure 1. In Table 1 the number of cattle tests and reactors that were carried out and disclosed, respectively, during the study period is presented. The highest percentage of reactors was disclosed in herds with dairy cattle only, or dairy and beef (i.e. young stock) (Table 2). One third of the study animals were tested once during the study period, 29% of the tests were carried out on animals under one year old, 25% on animals between one and two years old and 46% on cattle over two years old (Table 3). Table 4 presents the age at which animals were tested and became reactors, by purchased or born on the farm.

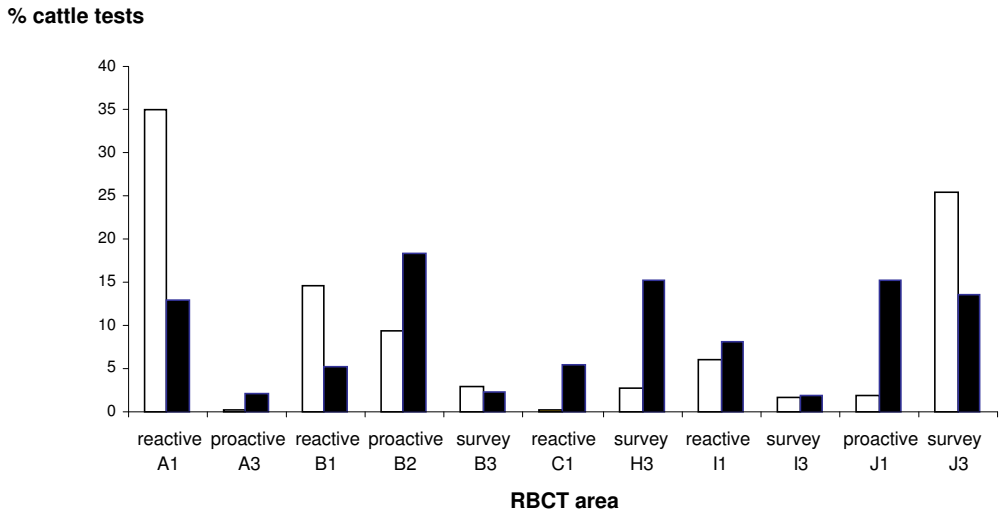


Fig.1 Percentage of cattle tests by restocking status and RBCT trial area.
Key: A1, A3: Gloucestershire and Herefordshire; B1, B2, B3: Devon and Cornwall; C1: East Cornwall; H3: Somerset and Devon; I1, I3: Gloucestershire; J1, J3: Devon, white bars are restocked farms, black bars continuously stocked farms.

Among the purchased cattle that had been exposed to bTB previously on the study farms there were 155 reactors/21,807 animal tests with unknown previous exposure to cattle with bTB on source farms, 56 reactors/9,792 animal tests not exposed to cattle with bTB on source farms and 32 reactors/ 3,866 animal tests exposed to cattle with bTB on source farms. Among purchased cattle that were not exposed on the study farm, 81 reactors/18,858 animal tests had unknown previous exposure and 21 reactors/4,973 animal tests were exposed to bTB on source farms. Only 39 reactors/12,033 animal tests, had not been exposed in the either study or source farms.

From Table 3 the risk of becoming a reactor increased with age and with the number of reactors to which an animal could have been potentially exposed. These are not independent risks; older cattle have more chance of being exposed.

Table 1. Distribution of 156,562 cattle tests in 697 herd tests and 723 reactors by test type

Test type	Number of cattle tests	% (cattle tests/ total)	Number of herd tests	Number of reactors	Reactors/ cattle tests x100
VE-SI	84,631	54.06	310	392	0.46
VE-CT	13,101	8.37	61	109	0.83
VE-6M	18,242	11.65	85	84	0.46
VE-CON	3,000	1.92	28	13	0.43
VE-CON6	5,236	3.34	32	28	0.53
VE-WHT	17,504	11.18	102	56	0.32
VE-RHT	1,442	0.92	8	2	0.14
VE-12M	4,378	2.80	23	15	0.34
VE-CON12	9,028	5.77	48	24	0.27
TOTAL	156,562	100.00	697	723	3.79

Table 2. Number and percentage of cattle tests and reactors from a total of 156,562 cattle tests and percentage of reactors out of the total 723 by herd purpose

Purpose of herd	Number of cattle tests	% of total cattle tests	Number of reactors	Reactor/ cattle tests x100	% of total reactors
Suckler only	6,810	4.35	12	0.18	1.66
Dairy only	51,046	32.60	272	0.53	37.62
Youngstock <30 m.o. only	19,598	12.52	47	0.24	6.50
Young stock and suckler	35,458	22.65	136	0.38	18.81
Young stock and dairy	43,650	27.88	256	0.59	35.41
TOTAL	156,562	100.00	723	0.46	100.00

Table 3. Total number of reactors per 1,000 cattle tests disclosed from 153,204 cattle tests and 697 reactors by age and number of previous exposures

Proportion of reactors (x1000 cattle tests) potentially exposed to by group	Age in years						
	≤1	>1-2	>2-3	>3-4	>4-5	>5-6	>6
None	0.6	1.9	6.4	8.8	13.0	7.8	10.5
One to five	0.0	1.6	4.8	11.8	12.9	10.1	8.7
Six to twenty	1.3	1.9	5.2	7.4	6.2	8.8	8.4
More than twenty	1.9	1.5	4.2	7.5	10.9	10.4	9.2

The number of reactors among cattle born on the study farms compared with those purchased was higher for all ages (Table 4).

Table 4. Total number of reactors per 1,000 cattle tests disclosed from 153,357 cattle tests and 697 reactors by age in years at the test and by birth location

Number of reactors (x1000 cattle tests) by birth location	Age in years						
	≤1	>1-2	>2-3	>3-4	>4-5	>5-6	>6
Born on study farm	0.6	1.8	4.0	9.9	12.0	10.0	11
Purchased	0.35	1.6	5.7	7.0	8.2	8.2	8.8

Multivariable analysis

Results from the multivariable analysis are presented in Table 5. There was a significant increased risk for a bovine being a reactor at a test when it had been exposed to a reactor at a previous test on the study farm. In addition, the OR was higher for cattle purchased than cattle born on the farm, although the CI overlapped for all three exposures. In a second model, the effect of cattle born or purchased (y/n) and previously exposed to reactors or not (y/n) was investigated using two variables. Both were associated with an increased risk of an animal being reactor at a test on the study farms. An interaction between the two variables indicated a further increased risk, suggesting that the exposure increased if the animal was purchased. The risk associated with restocking and the type of test was similar to that presented in Table 5.

With adjustment for previous exposure to bTB, cattle on restocked herds had a reduced risk of being a reactor. There were no significant effects of RBCT treatment although the OR were >1 for both reactive and proactive treatments. Cattle were less likely to be a reactor at a follow up test after the initial herd test.

Table 5. Multivariable multilevel logistic regression with random effects analysis from 156,562 cattle tests using 697 herd tests in 144 herds and 48,055 individual cattle

Explanatory variable	Level	Number reactors/ number animal tests	Coef.	OR	S.E.	95% CI
Restocking	continuously stocked	666/125,249	ref			
	restocked	57/31,313	-0.86	0.42	0.42	0.19 – 0.96
RBCT treatment	survey only	193/51,560	ref			
	reactive	270/56,956	0.54	1.72	0.39	0.80 – 3.69
	proactive	260/48,046	0.60	1.82	0.44	0.77 – 4.27
Test type	yearly	97/32,352	ref			
	short interval	392/84,631	-0.74	0.48	0.15	0.35 – 0.64
	other strategic*	234/39,579	0.29	1.34	0.15	1.00 – 1.80
Potential exposure to reactors prior to test	b** not exposed	61/28,548	ref			
	b** exposed	278/56,453	0.81	2.25	0.16	1.66 – 3.06
	p*** not exposed in study	141/35,864	1.13	3.08	0.17	2.23 – 4.26
	p*** exposed in study	243/35,465	1.32	3.74	0.16	2.72 – 5.13
	variance between herds		2.58		0.61	
	variance between animals within herds		0.52		0.20	

* includes tests other than short interval tests and yearly tests, b** born on study farm, p*** purchased onto study farm

The result for residuals at level 3 (variance between herds) was $\chi^2 = 5.89$, $p=0.75$ and at level 2 (cattle within herds) was $\chi^2 = 71.8$, $p<0.01$ which suggested that the model fit better at herd than animal within herds level.

DISCUSSION

Purchased cattle were at a higher risk of becoming reactors, whether they were exposed on the study farm or not (Table 5). This might indicate that these cattle brought infection with them onto the study farm or that these cattle were less immune to infection on the study farm than homebred cattle once they arrived. Both hypotheses are interesting and suggest that purchased cattle are increasing the occurrence of reactors. Individual factors such as age at test and dairy cattle have long been suggested to be at higher risk of becoming reactors at a test (Francis, 1947). The results presented here highlight the importance of the potential exposure to other infected cattle as source of infection as the main source of infection to other cattle. This was highly correlated with age. The purpose of the herd and age can be explained by these results and the current results are more useful because it is not possible to alter a bovine's age and difficult to change a farms purpose.

In the final model restocked herds versus continuously stocked and short interval tests versus other tests had a lower risk of HBD than herds. The most plausible explanation for the reduced risk for restocked herds is that cattle in newly formed herds are the main source of bTB because destocked farms have had a period of time with no cattle. Previous exposure to cattle with bTB was included in the model and consequently the residual risk was only the farm effect, this was lower for restocked farms than for continuously stocked farms. This is similar to the pattern observed by Carrique Mas et al. (2007) in a study of all restocked herds, where the time since the last bTB breakdown on the farm was an exponentially decaying risk for the new cattle tested at the first test after restocking. The lower risk for a bovine reacting at a short interval tests is likely to have a different explanation. A short interval test occurs approximately 60 days after detection of at least one reactor at a routine herd test. Consequently, recent reactors are removed and the SI test detects only reactor cattle that are either recently infected or that did not (for some reason) react at the routine herd test. (Given the sensitivity of the test, it may be that some cattle never react and so these cattle may not react to a second test, so the SI test may not remove all residual infection). All SI tests finish with a SI test with no reactors (two such tests if the HBD was confirmed (Green and Cornell, 2005)), this is probably why these tests are at apparent lower risk for disclosing a reactor.

The RBCT treatments were not associated with a significant risk of a bovine animal becoming reactor, however, this variable was forced in the model as it was one of the two variables used in the study design criteria. There were far fewer farms in the current study than in the RBCT and it is not surprising that statistical significance was low. It is of interest that the OR indicated a positive non-significant risk and a further study of all cattle prior exposure to bTB in the RBCT herds together with estimates of badger infection might be very revealing.

ACKNOWLEDGEMENTS

The study was funded by DEFRA (Project SE3026). The authors are grateful to Andy Mitchell from the VLA for providing the farms and VetNet data; to Alan Aldridge for the BCMS data and to farmers for their collaboration in the study.

REFERENCES

Bourne, J., Donnelly, C., Cox, D., Gettinby, G., McInerney, J., Morrison, I. and Woodroffe, R. (1999). An Epidemiological Investigation into Bovine Tuberculosis. Second Annual Report of the Independent Scientific Group on Cattle TB MAFF Publications, London

Carrique-Mas J.J., Medley, G.F., Green, L.E. (2007). Risks for bovine tuberculosis in British cattle farms restocked after the foot and mouth disease epidemic of 2001. Prev. Vet. Med. on line Dec 2007

Costello, E., Egan, J. W., Quigley, F. C. and O'Reilly, P. F. (1997). Performance of the sigle intradermal comparative tuberculin test in identifying cattle with tuberculous lesions in Irish herds. Vet. Rec. 141, 222-224

DEFRA (2005). State Veterinary Service. Procedures and Emergency Routines. Veterinary instructions ("VIPER"), Chapter 23, Section G

Green, L.E., Cornell, S.J., 2005. Investigations of cattle herd breakdowns with bovine tuberculosis in four counties of England and Wales using VETNET data. Prev. Vet. Med. 70, 293-311.

Francis, J. (1947). Bovine tuberculosis including a contrast with human tuberculosis. London, Staples Limited

Gibbens, J. C., Sharpe, C. E., Wilesmith, J. W., Mansley, L. M., Michalopoulou, E., Ryan, J. B. 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

Hosmer, D. W. and Lemeshow, S. (2000). Applied Logistic Regression, 2nd Edition. New York: Wiley

Krebs, J. R., Anderson, R., Clutton-Brock, T., Morrison, I., Young, D. and Donnelly, C. (1997). Bovine Tuberculosis in Cattle and Badgers Report. Ministry of Agriculture, Fisheries and Food, HMSO. London

Mitchell, A. P., Green, L. E., Clifton-Hadley, R., Mawdsley, J., Sayers, R., Medley, G. F. (2006). An analysis of single intradermal comparative cervical test (SICCT) coverage in the GB cattle population.Proceedings of the Society for Veterinary Epidemiology and Preventive Medicine (SVEPM) 70-86, Exeter, 29th – 31st March

Monaghan, M. L., Doherty, M. L., Collins, J. D., Kazda, J. F. and Quinn, P. J. (1994). The tuberculin test. Veterinary Microbiology 40, 11-124

Neill, S. D., O'Brien, J. J. and Hanna, J. (1991). A mathematical model for Mycobacterium bovis excretion from tuberculous cattle. Vet. Microbiol. 28, 103-109

Pollock, J. M., Rodgers, J. D., Welsh, M. D. and McNair, J. (2006). Pathogenesis of bovine tuberculosis: The role of experimental models of infection. Vet. Microbiol. 112, 141-150

Rasbash, J., Steele, F., Browne, W. and Prosser, B. (2004). A User's Guide to MLwiN. Version 2.0. , Centre for Multilevel Modelling, Institute of Education, University of London

BIOSECURITY

TYPE AND FREQUENCY OF CONTACTS BETWEEN BELGIAN PIG HERDS

S. RIBBENS*, J. DEWULF, F. KOENEN, K. MINTIENS, A. DE KRUIF AND D. MAES

SUMMARY

Knowledge of direct and indirect contacts between pig herds is a requirement for understanding the potential between-herd transmission of pathogens. Different contacts between Belgian pig herds were investigated. Data was obtained by a postal survey and by analysis of livestock movements. Direct contacts included transports of pigs by onto-farm, off-farm and between-farm movements. Indirect contacts included vehicles and visitors. The median number of direct contacts per herd made by onto-farm movements was 0.2/month (Q1: 0; Q3: 0.5). A zero-inflated negative binomial regression model was used to describe differences in the number of onto-farm movements according to herd size and herd type. The median number of origin herds during an 8-month period for between-farm movements was 4 (Q1: 2; Q3: 8). Directed graphs of between-farm movements were constructed. The median number of vehicles entering a herd and visitors entering the animal housing was 8.3/month (Q1:5.7; Q3:12.6) and 3.0/month (Q1: 1.5; Q3: 6).

INTRODUCTION

In Belgium, pig production is one of the most intensified sectors in livestock production. In 2005, the number of pigs in Belgium was approximately 6,300,000, divided among 8,510 pig herds (Sanitel-Pigs, 2005). The total surface of Belgium is 30,527.9 square kilometres, divided in 10 provinces and 589 municipalities (Belstat, Federal Public Service for Economy of Belgium). Regions with a high livestock density have been shown to be very vulnerable to disease outbreaks (Miry et al., 1991; Koenen et al., 1996; Elbers et al., 1999; Stegeman et al., 2004). Pig production is mainly concentrated in northern Belgium, resulting in several densely populated livestock areas (DPLAs) (Michel and Windhorst, 2003).

The intensity and frequency of different contacts between pig herds will to a large extent determine the spread of infection in different regions. This is especially the case during the period between the introduction of an epidemic and the first diagnosis of infection, also called the “high-risk period” (HRP) (Elbers et al., 1999). The importance of direct and indirect contacts for spreading of infections is clearly illustrated by the efficacy of movement restrictions in limiting outbreaks (Stegeman et al., 1999; Thrusfield et al., 2005). Spreading of an infection either in space or time depends both on infection specific parameters (e.g. stability of the pathogen, amount excreted, etc.) and the type and frequency of contacts (Fèvre et al., 2006).

The type of contact can be divided into direct (from animal to animal) or indirect (using one or more intermediate steps) (Ribbens et al., 2004). Each type comprises a certain risk for

*Stefaan Ribbens, University of Ghent, Faculty of Veterinary Medicine, department of Fertility, Obstetrics and Herd Health, Merelbeke, Belgium. Email: Stefaan.Ribbens@UGent.be

transmission of an infectious agent which again varies for different infections (Amass and Baysinger, 2006). Undoubtedly, for most infections, the highest risk is through direct animal contact. In daily management, this can occur through animal movements when purchasing animals and through shows or meetings. Indirect types of contact with less obvious, although not negligible contact include vehicles, persons and other animals (pets, rodents) entering the premises (Ribbens et al., 2004). During epidemics in DPLAs, the so called 'neighbourhood infections' have been important (Crauwels et al., 2003; Mintiens et al., 2003). Transmission is assigned to this route when spread of infection occurs between herds located in close vicinity (less than 1 km) and when no specific transmission route can be identified.

Accurate knowledge on the types and frequency of contacts between herds is both important for tracing purposes during contagious animal-disease outbreaks and for the understanding and prediction of the impact of particular infection-control strategies (Morris et al., 2002). A number of surveys have already been conducted to quantify different types of contacts between livestock farms (Sanson et al., 1993; Sanson, 2005; Nielen et al., 1996; Bates et al., 2001), but only one that was specific for pigs (Stärk, 1998). Available studies showed that contact structures vary in space and/or time. Papers using identification-and-registration (I&R) data on pig movements are rather scarce (Bigras-Poulin et al., 2007). Without accurate knowledge of the contact structure, infection models can have difficulties predicting the potential size of an outbreak (Kao, 2002; Green et al., 2006).

Today, little is known about the type and frequency of contacts between Belgian pig herds. The aim of this paper was to describe and quantify the different types of contact by using information from a postal survey and from available I&R data on livestock movements.

MATERIALS AND METHODS

Data Sources

Identification and registration data

The data concerning animal movements were available from the I&R database for all Belgian pig herds for an 18-months period (2005-2006). The information available comprised all recorded onto- and off-movements (i.e. inbound and outbound movements) of pigs for each herd in Belgium, including the number of animals transported. Recording pig movements is legally required in Belgium (Ammendrup and Barcos, 2006). The majority of I&R data on livestock movements were recorded using a computerized system, only a minority of smaller hauliers use written records for the registration of animal movements (Sanitel-pigs, personal communication). A movement is related to a geographical location with proper x- and y-coordinates, so separated sites of the same farm are considered separate farms. For a subset of the database (8 last months) also the individual identification of the transport vehicle and the type of animals transported were available. This additional information made it possible to link the onto- and off-movements on different herds for this time period. Sequential movements (i.e. the same transport truck having several stops to load or unload animals) were not traced in the dataset with linked movements. Based on available x and y coordinates of the registered herds, the Euclidian distances (i.e. straight-line distances) for between-farm movements were calculated using Pythagorus' theorem.

Movements with incomplete data and herds with incomplete coordinates were removed from the dataset before analysis. The locations of the movements were linked to a dataset with

all active Belgian pig locations (including herd size and herd type). Movements towards pig traders or movements towards locations other than a pig herd or the slaughterhouse were removed before analysis. From all the herds present in the I&R database, only those that had pigs counted during the latest available report of the obligatory tri-annual visit by the herd veterinarian, were considered active. Onto-farm movements did include animal imports from foreign countries and separate data on animal imports recorded in TRACES were used (Kroschewski et al., 2006). Data on between-farm movements did not include pig movements from abroad, but only movements between two Belgian pig herds.

The number of pigs per square km was calculated per municipality. For the surface areas of each Belgian municipality, we consulted Belstat (Federal Public Service for Economy of Belgium). A threshold value for DPLAs, as defined by Michel and Windhorst (2003) namely, areas are dense in terms of risk for classical swine fever (CSF) when over 300 pigs are counted per square km, was used. Below this value, areas were defined as sparsely populated livestock areas (SPLAs).

Written questionnaire data

The administration and design of the questionnaire is described in detail in Ribbens et al. (2008). In summary, a questionnaire was distributed by mail in 2005 to 609 pig herds (stratified random sample). It covered several aspects regarding biosecurity, management practices and different types of contacts between farms. The questionnaire was made in two languages (Flanders= Dutch, Wallonia= French) and checked for consistency by a bilingual researcher. Before administration, the form was pre-tested during a herd visit on 7 pig farms. The questionnaire was semi-closed, pre-coded and is available upon request to the first author (in Dutch and French). Time taken to fill in the questionnaire was approximately 10 minutes. Issuing a question twice assessed the repeatability; this was done for two topics. Validity was evaluated by comparing data provided in the questionnaire with information available in the I&R database.

Questions comprised information on direct contacts, such as whether piglets or breeding replacement stock were purchased including details on the annual number of animals bought, the frequency of purchase, origin of the animals, manners of transport and general biosecurity aspects. Questions concerning indirect contacts such as an estimation of the number of vehicles entering the herd and the number of professional and non-professional visitors with potential animal contact (i.e. entering the herd's stables) were also included.

Data presentation and analysis

Descriptive results (minimum, 25th percentile, median, 75th percentile and maximum) were generated and were generally presented per one month period. This time period was deliberately chosen since this corresponds relatively well to the length of a high-risk period (HRP) observed for some of the most important epidemics (e.g. CSF) in the recent past (Elbers et al., 1999). The duration of the HRP largely determines the magnitude of an outbreak as the infection may circulate freely during this time (Horst et al., 1998). For different types of contact and herd size, the Spearman's rank correlation coefficient was calculated and different non-parametric techniques were used as the frequency of occurrence of all types of contacts deviated from the normal distribution (Mann-Whitney test, Kruskal-Wallis Test, Wilcoxon Signed Ranks test).

To investigate the relationships between the number of onto-farm movements and potential predictor variables (herd size and herd type), a zero-inflated negative binomial regression model

(ZINB) was conducted. Poisson regression models are often used to describe count data (Neter et al., 1996). These assume that mean and variance are the same, which is rarely the case. Negative binomial models are a solution to correct for this extra-Poisson variation or over-dispersion (Hilbe, 2007). When a high proportion of zeros is present, zero-inflated models are appropriate (Long, 1996). These models can deal with additional over-dispersion caused by an excess of zeros, this through a splitting process that models the probability of a zero outcome by logistic regression, while the count outcome is modelled using a Poisson or negative binomial structure. We used the likelihood ratio test and the Vuong statistic as indicators of goodness of fit (Dohoo et al., 2003). For evaluating the over-dispersion, we performed a likelihood ratio test which compares a Poisson model to the negative binomial model. The Vuong statistic was used to decide if a zero-inflated model was better suited compared with regular Poisson or negative binomial models.

Statistical analyses were performed using SPSS 14.0 (SPSS Inc., Chicago, Illinois) except for ZINB (Stata 9 (Strata Corporation, College Station, Texas, USA)). The map of the location of the registered pig herds was constructed using ArcMAP 9.2 (ESRI, Redlands, CA, USA). To demonstrate relationships through between-farm movements, directed graphs were constructed using the Kamada-Kawai algorithm in the network software Pajek (2003) (DiBattista et al., 1999). Farms are represented by a circle ('vertices'), which are connected through weighted lines ('edges'). The arrows give the direction of the movement. Directed networks were constructed for a one month time period (April 2006). The Kamada-Kawai graph layout attempts to position vertices on the space so that the geometric (Euclidean) distance between them is as close as possible to the graph-theoretic (path) distance between them (Kamada and Kawai, 1989).

RESULTS

The dataset of active Belgian pig herds consisted of 8,510 herds. Locations were considered inactive due to lack of a recent visit report (1,346 herds) or absence of pigs (436 herds); 406 herds were left out because of missing data (herd size or coordinates). During the 10-month period in 2005, 6,400 registered locations made in total 48,655 onto-movements. Of these locations, 631 could not be linked with the dataset of 'active' Belgian pig herds. Fifty-five of these locations were pig traders, justifying for 21.4% of onto- movements (10,412) made during the 10-month time-frame. This left 576 herds that could not be linked (accounting for 5.9% (2,862) of made onto-movements). Linkage failed because these herds were considered inactive or had missing data (coordinates or herd size).

Response for the questionnaire was 436/609 (71.6%) with 421 respondents suitable for analysis. Responding herds are plotted in Fig. 1.

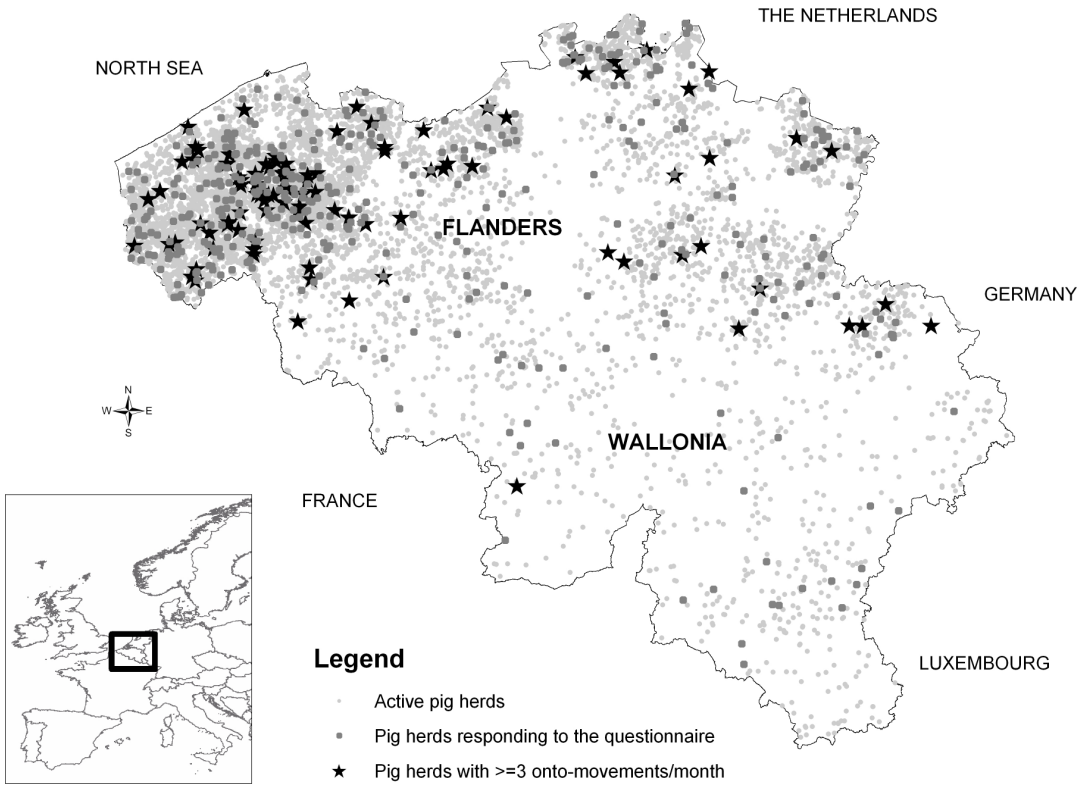


Fig. 1 Geographical location of active pig herds in Belgium (N=8,510), pig herds responding to the questionnaire (n=421) and pig herds with 3 or more onto-farm movements per month (n=112).

General description of the Belgian pig industry

Data on herd characteristics of Belgian pig herds can be found in Table 1. The majority of pig farms were situated in the Northern part of the country (Fig. 1). Of all municipalities in Belgium, about 18.7% had ≥300 pigs/km². These municipalities with DPLAs accounted for 85.1% of the total pig population. Within the municipalities, pig herds were geographically clustered as the majority of responding farmers (59.6%) reported having other pig herds within a 500 metre radius; another 27.6% reported having other pig herds within a 1 km radius.

Table 1. Descriptives of Belgian pig industry.

VARIABLE	N [a]	MIN [b]	Q1 [b]	Q2 [b]	Q3 [b]	MAX [b]
Pig density in Belgian municipalities [c]	589	0	0.5	25.5	164.0	3,185.8
Number of animals in Belgian pig herds						
- boars	8,510	0	0	1	1	145
- finishers & future replacement stock	8,510	0	21	348	725	9,200
- sows	8,510	0	0	20	105	3,350
- weaner piglets	8,510	0	0	30	280	8,722
TOTAL PIGS	8,510	1	160	502	1,046	11,919

[a]: number of municipalities/pig herds; [b]: Minimum, 25th percentile, 50th percentile, 75th percentile and maximum values of the variable; [c]: calculated densities.

Direct Contacts

Onto-farm movements

The median number of onto-farm movements per month irrespective of herd type was 0.2 (Table 2). The distribution is highly skewed, with 2,741 (32.2%) of the herds with no onto-movements and another 2,550 (30.0%) herds with less than 0.3 onto-movements/month. A small number of pig herds (1.2% or 104 herds) had a high number of onto-farm movements per month (3 or more onto-farm movements per month, see Figure 1). The geographical distribution of the herds that had 3 or more onto-farm movements/ month is presented in Fig. 1. These herds were predominantly herds purchasing piglets (fattening herds (71%) or piglet multipliers (29%)).

Table 2. Onto-farm movements in Belgian pig herds.

ONTO-FARM MOVEMENTS PER MONTH [a]	N [b]	MIN [c]	Q1 [c]	Q2 [c]	Q3 [c]	MAX [c]
- Breeding herd	736	0	0	0.1	0.5	4.2
- Farrow-to-finish herd	4,690	0	0	0.1	0.4	3.4
- Finishing herd	2,959	0	0.1	0.4	0.8	9.8
- Piglet multiplier	125	0	0.5	1.1	2.2	13.4
TOTAL HERDS	8,510	0	0	0.2	0.5	13.4

[a]: calculated with I&R data for a 10-month period in 2005; [b]: number of pig herds; [c]: Minimum, 25th percentile, 50th percentile, 75th percentile and maximum values of the variable.

As the number of onto-farm movements was count data with many zero-valued observations, a zero-inflated negative binomial (ZINB) regression model was indicated. The variance:mean ratio was 11.01 (45.80:4.16), indicating over-dispersion. The likelihood ratio test was highly significant ($P<0.001$) and the value of α (1.11; CI: 1.02-1.20) suggested that a negative binomial model was preferable to a Poisson model as ($1+\alpha*$mean): $(1+(1.11*4.16))=5.60$ (Dohoo et al., 2003)). The Vuong test ($Z=6.13$; $P<0.0001$) suggested a zero-inflated model had a superior fit (Dohoo et al., 2003). Results of the ZINB model can be found in Table 3.

The negative binomial portion of the regression model (Table 3) showed that frequency of movements onto farms differed significantly between breeding herds and every other categories of herd; in the logistic portion of the model finishing herds and piglet multiplier herds were

significantly different to breeding herds. Coefficients for onto-farm movement of the negative binomial portion of the model for piglet multiplier was 1.5 and finishing herds 0.48 compared with the baseline (breeding herds). Farrow-to-finishing herds were predicted to have the least movements. Although a significant predictor, a one point increase in herd size, would only lead in an expected increase in number of onto-farm movement with a factor of exp (0.0002) while holding all other variables in the model constant. In the logistic portion of the model a negative regression coefficient indicates a higher likelihood of onto-farm movements.

Table 3. Coefficients' estimates of variables included in a zero-inflated negative binomial regression model for onto-farm movements (calculated with I&R data for a 10-month period in 2005).

VARIABLE	COEFF. [a]	SE [b]	P	LOW 95% CI	HIGH 95% CI
Negative-binomial portion					
Herd size	0.0002	0.0000	0.000	0.0001	0.0003
Herd type:					
- breeding herd	Baseline				
- farrow-to-finishing herd	-0.2722	0.0583	0.000	-0.3864	-0.1579
- finishing herd	0.4891	0.0584	0.000	0.3547	0.5836
- piglet multiplier	1.5087	0.1153	0.000	1.2828	1.7346
Constant	1.3619	0.0578	0.000	1.2486	1.4752
Logistic portion (probability of zero count)					
Herd type					
- breeding herd	Baseline				
- farrow-to-finishing herd	-0.0247	0.1647	0.881	-0.3474	0.2981
- finishing herd	-1.9893	0.4008	0.000	-2.7749	-1.2037
- piglet multiplier	-1.7247	0.7304	0.018	-3.1562	-0.2931
Constant	-1.3625	0.1624	0.000	-1.6809	-1.0411
A	1.1068	0.0478		1.0167	1.2046

[a]: Coefficient; [b]: standard error; Number of observations=8,510; Nonzero observations=5,760; Zero observations=2,741; Likelihood ratio test chi² (d.f.=4)=839.39; *P*=0.0000; Vuong test Z=6.13, P=0.0001).

Off-farm movements to the slaughterhouse

On average 14,279 slaughterhouse movements per month were made in Belgium with, on average, 753,328 animals transported per month. The off-farm movements towards the slaughterhouse reflect the general structure of the Belgian pig industry. The majority of these movements occurred within the same province (64%). Four slaughterhouses in Western-Flanders account for one-third of all pigs slaughtered in Belgium.

Animal imports

Of the respondents purchasing piglets, 13.5% purchase from neighbouring countries (Table 4). In 2006, on average 100,446.4 animals were imported each month in 644.3 import movements/month (TRACES, 2006). About 46% were slaughter pigs; about 41% breeding stock and about 13% were fatteners.

Table 4. Descriptive statistics on direct contacts in Belgian pig herds.

DIRECT CONTACTS [a]	N [b]	MIN [c]	Q1 [c]	Q2 [c]	Q3 [c]	MAX [c]
Purchase of piglets	156					
- number of piglets/year	156	1	500	1,200	2,185	15,000
- number of purchases/year	156	1	2.00	2.66	12.0	52
- number of supplier farms/year	156	1	1	1	4	50
Purchase of piglets from abroad	21					
Piglets of unknown health status	17					
Purchase of replacement stock	185					
- number of replacement stock/year	185	1	5	30	60	400
Replacement stock of high health status	76					
Replacement stock from abroad	29					

[a]: calculated with survey data (August 2005); [b]: number of pig herds; [c]: Minimum, 25th percentile, 50th percentile, 75th percentile and maximum values of the variable.

Between-farm movements

Frequency and contact structure of between-farm movements:

In total 33,234 between-farm movements were recorded during the 8-month period. This corresponds with an average of 4,154.3 between-farm movements/month for the Belgian pig population. About 11.7% of all transported animals were replacement stock (adult boars and sows or non-adult replacement stock) and the majority (88.3%) were piglets or future finishing pigs (Table 5).

Table 5. Between-farm movements in Belgian pig herds.

BETWEEN FARM MOVEMENTS [a]	N [b]	MIN [b]	Q1 [b]	Q2 [b]	Q3 [b]	MAX [b]
Number of origin herds	33,234	1	2	4	8	181
Number of animals transported						
- breeding animals	3,884	1	5	8	10	300
- piglets	29,350	1	30	63	120	999
Distance of between farm movements (km)	33,234	0	8.1	19	36	250.5

[a]: calculated with I&R data for a 8-month time period in 2006; [b]: Minimum, 25th percentile, 50th percentile, 75th percentile and maximum values of the variable.

Between-farm movements made in April 2006 were plotted using directed graphs. Figure 2 shows between-farm movements through piglet transports. A selection of 158 pig herds, receiving at least 5 piglet movements was made. These herds received piglets from 929 different source herds.

Fig. 2 Directed network of 1,493 piglet movements between 1,078 Belgian pig herds in April 2006 (929 source herds, 158 recipient herds with at least 5 purchases a month).

In Fig. 3, distribution movements of replacement stock are illustrated; 128 pig herds distributed future replacement stock towards 465 recipient herds.

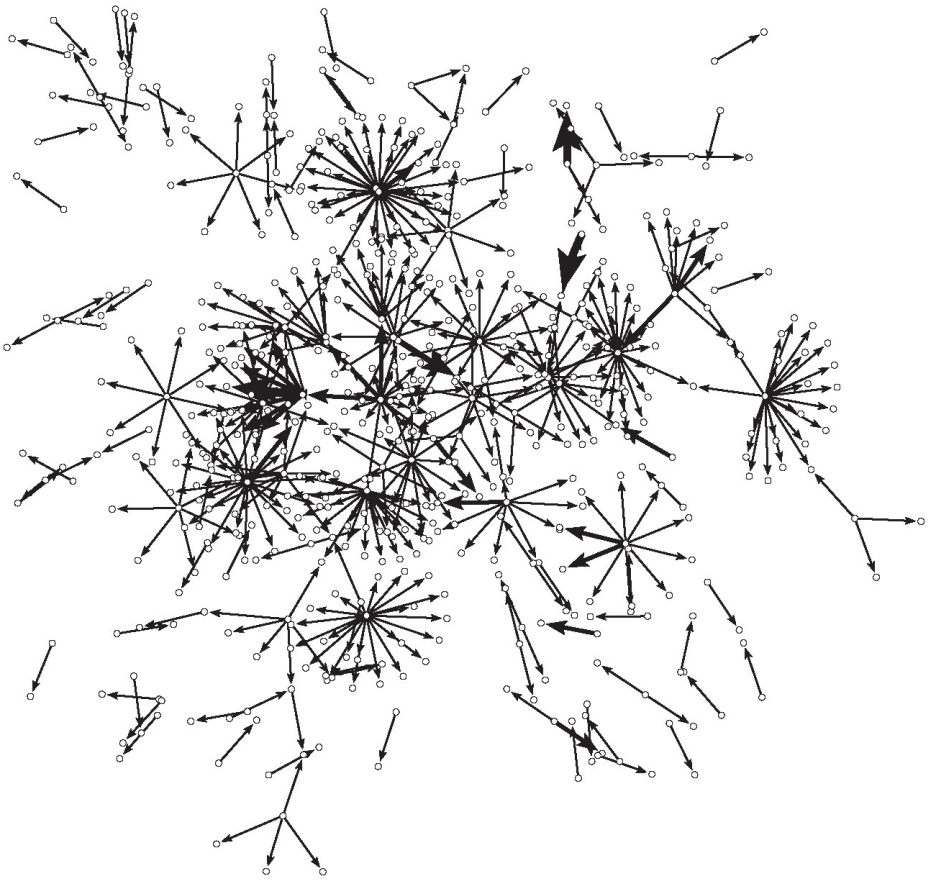


Fig. 3 Directed network of 520 replacement stock movements between 570 Belgian pig herds in April 2006 (128 source herds, 465 recipient herds).

For a one year time period, the median number of origin herds (regardless the type of pigs transported) was 4, with a fifth of the farms (19.2%) receiving animals from more than 10 different sources (Table 5). The number of different suppliers estimated on the basis of the I&R data was a little higher than the estimate made by farmers responding to the questionnaire (Table 4). Almost half (49.4%) of the farmers that were buying piglets and 17.8% of the farms purchasing replacement stock had more than 1 supplier farm.

Distances and location of between-farm movements

The median distance of between-farm movements was 19.0 km (Table 5). The majority of between-farm movements occurred within DPLAs (75.7%) and only 6.6% within SPLAs. Movements also happened between different regions: 10.9% go from DPLA to SPLA and 6.8% the other way round. Five percent of all between-farm movements were over 100 km. There was no significant difference in the distances travelled by replacement stock (median 26.7 km) or piglets/finishers (median 29.2 km) (Mann-Whitney test, P=0.531).

Indirect Contacts

Movement of livestock trucks and other trucks entering the premise

Feed trucks are the most common vehicle type entering pig farms (Table 6). The correlation between the number of vehicles that entered per month and the herd size was 0.745 (Spearman's ρ, P<0.01). Also, significantly more vehicles entered herds in DPLAs compared to SPLAs (Mann-Whitney test, P<0.001). In mixed herds, more vehicles entered compared to herds with only finishers or only sows (Kruskal-Wallis test, P<0.001) as for professional pig herds compared to hobby herds (Mann-Whitney test, P<0.001).

Table 6. Descriptive statistics on indirect contacts made by vehicles in Belgian pig herds.

VEHICLES ENTERING THE HERD [a]	N [b]	MIN [c]	Q1 [c]	Q2 [c]	Q3 [c]	MAX [c]
Vehicles entering the herd						
- livestock trucks/month	403	0	1.00	2.00	4.00	31
- trucks from feeding company/month	415	0	2.00	3.50	4.00	100
- trucks from rendering plant/month	398	0	1.00	2.00	4.00	25
-slurry trucks/month	404	0	0.00	0.17	0.83	16.67
Total vehicles/month	378	0	5.77	8.33	12.66	131

[a]: calculated with survey data (August 2005); [b]: number of pig herds; [c]: Minimum, 25th percentile, 50th percentile, 75th percentile and maximum values of the variable.

Professional and non-professional visitors entering the animal facilities

The veterinarian was the frequent professional visitor entering the animal housing (Table 3). In general, significantly more professional visitors entered the pig herds compared to non-professional visitors (P<0.001). In professional and especially non-professional herds, a large variation existed in the number of persons entering. Spearman's ρ for the number of persons entering and herd size was 0.209 (P<0.01) for professional and -0.086 (P=0.078) for non-professional visitors.

Non-professional visitors entered herds significantly more frequently in SPFA's compared with DPLA's (P=0.002). More professional visitors entered professional pig herds (P<0.001) and more non-professional visitors entered hobby herds (P<0.001).

Table 7. Descriptive statistics on indirect contacts made by visitors with potential animal contact in Belgian pig herds.

VISITORS WITH POTENTIAL ANIMAL CONTACT [a]	N [b]	MIN [c]	Q1 [c]	Q2 [c]	Q3 [c]	MAX [c]
Professional visitors with potential animal contact						
-other swine farmer/month	417	0.00	0.00	0.00	0.00	20
-veterinarian/month	411	1.00	1.00	1.00	2.00	15
-advisor/month	416	0.00	0.00	0.00	1.00	8
-climatologist/month	417	0.00	0.00	0.00	0.00	2
-animal transporter/month	417	0.00	0.00	0.00	0.50	8
-pig trader/month	414	0.00	0.00	0.00	0.18	5
Total professional visitors/month	410	0.00	1.00	2.12	4.50	35
Non-professional visitors with potential animal contact						
- family/month	416	0.00	0.00	0.00	0.00	60
- neighbours/month	417	0.00	0.00	0.00	0.00	20
-friends/month	417	0.00	0.00	0.00	0.00	6
Total non-professional visitors/month	416	0.00	0.00	0.00	0.50	60
Total visitors/month	409	0.00	1.50	3.00	6.00	64

[a]: calculated with survey data (August 2005); [b]: number of pig herds; [c]: Minimum, 25th percentile, 50th percentile, 75th percentile and maximum values of the variable.

DISCUSSION

In previous studies that aimed at describing different types of contacts, data were mostly collected in a relative small geographically-defined area using on farm questionnaires in which farmers were asked to record all direct and indirect contacts during a certain time frame (Sanson et al., 1993; Nielen et al., 1996; Stärk, 1998; Bates et al., 2001). A potential drawback of this methodology is that the numbers of contacts are estimates given by respondents, and are not necessarily representative for the pig industry in the whole country. Nowadays, I&R data record livestock movements for the entire country. They are helpful gaining insight on the contact structure of the pig industry as illustrated by Bigras-Poulin et al. (2007). I&R data are not, for instance, subject to recall bias. Still, the use of survey data was indispensable for estimating the indirect contacts, and these contacts may contribute to infection spread.

Direct animal contacts are undoubtedly most hazardous for transmission of infection. Undetected infected animals can transport infection between herds. Analysis of I&R data revealed a certain percentage of pig herds made a high number of direct contacts. Because of the frequent purchase of animals on these herds, they are at relatively high risk since they had a higher probability of introducing infections through animal movements. All high contact herds (at least 3 onto-movements/month) are situated in the DPLAs (Fig. 1), where they can easily become the source for spread of infection through neighbourhood or/and other contacts.

The number of different suppliers estimated on the basis of the I&R data was a little higher than the estimate made by farmers responding to the questionnaire (Table 4). This may indicate that the farmers are not always fully aware of the large number of different suppliers they buy from. In a study by Maes et al. (2004) mixing of animals from different origin herds was found to be a significant risk factor for mortality in finishers. Also for epidemic infections, mixing of animals is a clear risk (Gibbens et al., 2001). Besides the frequency of contacts, also the knowledge of the sanitary status of the origin farm is crucial for prevention of infections. Nevertheless, 10.9% of the respondents claim they do not have prior knowledge concerning the health status of the origin herd (Table 4). Demands regarding the sanitary status of the purchased animals are naturally higher for replacement stock, although about 30% never uses a quarantine period for its replacement stock (Ribbens et al., 2008). Live animals are imported every year into Belgium. This increases the risk of introduction of foreign infections (Miry et al., 1991; Koenen et al., 1996; Bouma et al., 2003), although the number of piglet imports has diminished in Belgium last years.

The structure of contacts determines the potential impact of infection spread. Graph drawings and networks make it possible to visualise the structure of the contacts. The 'spider'-structure of between-farm movements can be observed in Figure 2 and Figure 3. In Figure 2, piglet movements showed a typical structure of the pig industry, with piglet producers distributing towards a recipient herd making these herds some kind of 'collectors' of infection. In Figure 3, distribution movements of replacement stock are illustrated; some of these herds can potentially spread disease to a large number of recipients. If disease remains undetected in these 'spreader-herds', infection can disseminate towards several farms because different movements happen in a short time-frame.

The distances travelled between herds may determine the extent infection can disperse. Although animal movement is in theory not distance dependent since trucks can easily drive a long distance, it was found that about a third (29%) of the between-farm movements were within a radius of less than 10 km. This is probably due to the high concentration of pig production within the DPLAs. Compared to other direct contacts, slaughterhouse movements incorporate very limited risks for transmission of infection if appropriate biosecurity measures are taken.

In terms of transmission of infection, indirect contacts are less efficient compared with direct animal-animal contacts (Amass and Baysinger, 2006). On the other hand, there are a large number of different indirect contacts that occur at a relative high frequency. A low probability multiplied with a high frequency may result in a medium risk. In the recent past, indirect contacts have been held responsible for spread of infection in several epidemics (Elbers et al., 1999; Gibbens et al., 2001).

The risk of transmission of infection depends on the type of vehicle: trucks coming from a rendering plant are generally considered to be a very high biosecurity risk. This risk is lower for empty livestock truck and trucks from the feeding company. As (empty) livestock and other trucks visit several herds on the same day, it is crucial they do not come into close or direct contact with the livestock present. As an illustration, farms with parking for pig transport vehicles located within 300 meters of the farm site were 9.28 times more likely to become reinfected with *Mycoplasma hyopneumoniae* or *Actinobacillus pleuropneumoniae* than farms with no parking site near the farm (Hege et al., 2002).

The results on indirect contacts made by visitors indicate that the number of persons coming into contact with the animals is only marginally correlated with the herd size, indicating that

owners of larger herd are probably more aware of the risks and limit the number of persons that may enter the herd to what is strictly necessary. This is also in agreement with the observation that non-professional visitors (friends, family, neighbours) enter more frequently on smaller herds which are often hobby herds. This difference in risk behaviour between professional and non-professional is in agreement with the results on biosecurity described in Ribbens et al. (2008).

Besides the known direct and indirect contact between pig herds, in every epidemic a fraction of the transmission of infection between herds remains unexplained. This is especially the case in regions with high pig densities where herds are located in close vicinity to each other. In Belgium, several DPLAs exist. It has been illustrated very clearly in previous CSF epidemics that neighbourhood transmission of infection occurred very often in these regions (Roberts, 1995). Although the through mechanism of these neighbourhood infections is still debated, several routes have been proposed (Ribbens et al., 2004).

There existed a large variation in the number of direct and indirect contacts of different Belgian pig herds. This study demonstrated a limited number of pig herds in Belgium have a very high number of direct contacts, sometimes with a large number of different herds.

Comparing the results of the current study with previously performed studies is difficult because of several reasons. All studies are performed on different populations with different herd sizes, herd locations and general herd structures. Moreover, contacts between livestock herds are time and area dependent and especially after major outbreaks, the structure of movements may heavily alter due to habit or legislation changes.

ACKNOWLEDGEMENTS

This research was supported by the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen: grant No. 41258). Sanitel-Pigs and Dr. J.M. Robijns are acknowledged for providing the necessary I&R data.

REFERENCES

Ammendrup, S. and Barcos, L.O. (2006). The implementation of traceability systems. Rev. sci. techn. Off. Int. Epiz. 25, 763-773.

Amass, S.F. and Baysinger, A. (2006). Swine disease transmission and prevention. In: Straw, B.E., Zimmerman, J.J., D'Allaire, S. and Taylor, D.J. (Eds.), Diseases of swine, 9th edition, Blackwell Publishing, UK, pp.1075-1098.

Bates, T.W., Thurmond, M.C. and Carpenter, T.E. (2001). Direct and indirect contact rates among beef, dairy, goat, sheep, and swine herds in three California counties, with reference to control potential foot-and-mouth disease transmission. Am. J. Vet. Res. 62, 1121-1129.

Belstat (Federal Public Service for Economy of Belgium). Surfaces of municipalities in Belgium, available at http://www.statbel.fgov.be (last consulted on 13 December 2006).

Bigras-Poulin, M., Barfod, K., Mortensen, S. and Greiner, M. (2007). Relationship of trade patterns of the Danish swine industry animal movements network to potential disease spread. Prev. Vet. Med. 80, 143-165.

Bouma, A., Elbers, A.R., Dekker, A., de Koeijer, A., Bartels, C., Vellema, P., van der Wal, P., van Rooij, E.M., Pluimers, F.H. and de Jong, M.C. (2003). The foot-and-mouth disease epidemic in the Netherlands in 2001. Prev. Vet. Med. 57, 155-166.

Crauwels, A.P., Nielen, M., Elbers, A.R., Stegeman, J.A. and Tielen, M.J. (2003). Neighbourhood infections of classical swine fever during the 1997-1998 epidemic in The Netherlands. Prev. Vet. Med. 61, 263-77.

DiBatista, G., Eades, P., Tamassia, R. and Tollis, I. (1999). Graph Drawing. Prentice Hall.

Dohoo, I, Martin, W. and Stryhn, H. (2003). Chapter 18: Modelling count and rate data. In: Veterinary Epidemiologic Research, Atlantic Veterinary College, Charlottetown, PE, Canada, pp. 391- 407.

Elbers, A.R., Stegeman, A., Moser, H., Ekker, H.M., Smak, J.A. and Pluimers, F.H. (1999). The classical swine fever epidemic 1997-1998 in the Netherlands: descriptive epidemiology. Prev. Vet. Med. 42, 157-184.

Fèvre, E.M., Bronsvoort, B.M., Hamilton, K.A. and Cleaveland, S. (2006). Animal movements and the spread of infectious diseases. Trends Microbiol. 14, 125-131.

Gibbens, J.C., Sharpe, C.E., Wilesmith, J.W., Mansley, L.M., Michalopoulou, E., Ryan, J.B. 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.

Green, D.M, Kiss, I.Z. and Kao, R.R. (2006). Modelling the initial spread of foot-and-mouth disease through animal movements. Proc. Biol. Sci. 273, 2729-2735.

Hege, R., Zimmermann, W., Scheidegger, R. and Stärk, K.D. (2002). Incidence of reinfections with Mycoplasma hyopneumoniae and Actinobacillus pleuropneumoniae in pig farms located in respiratory-disease-free regions of Switzerland – identification and quantification of risk factors. Acta Vet. Scand. 43, 145-156.

Hilbe, J.M. (2007). Overview of count response models (Chapter 1). In: Negative Binomial Regression, Cambridge University Press, pp. 8-18.

Horst, H.S., Dijkhuizen, A.A., Huirne, R.B.M. and De Leeuw, P.M. (1998). Introduction of contagious animal diseases into The Netherlands: elicitation of expert opinions. Livest. Prod. Sci. 53, 253-264.

Kamada, T. and Kawai, S. (1989). An algorithm for drawing general undirected graphs. Information processing letters 31, 7-15.

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.

Koenen, F., Van Caenegem, G., Vermeersch, J.P., Vandenheede, J. and Deluyker, H. (1996). Epidemiological characteristics of an outbreak of classical swine fever in area of high pig density. Vet. Rec. 139, 367-371.

Kroschewski, K., Kramer, M., Micklich, A., Staubach, C., Carmanns, R. and Conraths, F.J. (2006). Animal disease outbreak control: the use of crisis management tools. Rev. sci. tech. Off. Int. Epiz. 25, 211-221.

Maes, D., Duchateau, L., Larriestra, A., Deen, J., Morrison, R.B. and de Kruif, A. (2004). Risk factors for mortality in grow-finishing pigs in Belgium. J. Vet. Med. Series B 51, 321-326.

Michel, I. and Windhorst, H.W. (2003). Densely populated livestock areas: Definition and spatial distribution in the European Union. A case study for five member states. In: Huirne, R.B, Windhorst, H.W. (Eds.), Development of prevention and control strategies to address animal health and related problems in densely populated livestock areas of the community, Final Report FAIR5-CT97-3566, European Commision, pp. 10-25.

Mintiens, K., Laevens, H., Dewulf, J., Boelaert, F., Verloo, D. and Koenen, F. (2003). Risk analysis of the spread of classical swine fever virus through 'neighbourhood infections' for different regions in Belgium. Prev. Vet. Med. 60, 27-36.

Miry, C., Castryck, F., Koenen, F., Broes, A. and Segers, E. (1991). Quelques aspects de l'epizootie de peste porcine classique en Belgique en 1990. Epidémiol. Santé Animal 20, 23-32.

Morris, R.S., Sanson, R.L., Stern, M.W., Stevenson, M. and Wilesmith, J.W. (2002). Decision-support tools for foot and mouth disease control. Rev. Sci. Tech. 21, 557-567.

Neter, J., Kutner, M.H., Nachtsheim, C.J. and Wasserman, W. (1996). Logistic Regression, Poisson Regression and Generalized Linear Models (Chapter 14). In: Applied linear statistical models, Fourth Edition, pp. 567-628.

Nielen, M., Jalvingh, A.W., Horst, H.S., Dijkhuizen, A.A., Maurice, H., Schut, B.H., van Wuijckhuise, L.A. and de Jong, M.F. (1996). Quantification of contacts between Dutch farms to assess the potential risk of foot-and-mouth disease spread. Prev. Vet. Med., 28, 143-158.

Pajek (2003). PAJEK, program for large network analysis. University of Ljubljana, Slovenia, http://vlado.fmf.unilj.si/pub/networks/pajek

Ribbens, S., Dewulf, J., Koenen, F., Laevens, H. and de Kruif, A. (2004). Transmission of classical swine fever. A review. Vet. Q. 26, 146-155.

Ribbens, S., Dewulf, J., Koenen, F., Mintiens, K., De Sadeleer, L., de Kruif, A. and Maes, D. (2008). A survey on biosecurity and management practices in Belgian pig herds. Prev. Vet. Med. 83, 228-241.

Roberts, M. (1995). Evaluation of the optimal size of restriction zones in disease control with particular reference to classical swine fever. In: Goodall, E.A. (Eds.), Annual Conference of the Society for Veterinary Epidemiology and Preventive Medicine. SVEPM Reading, UK, pp. 119-130.

Sanitel-Pigs (2005 & 2006). Identification-and-registration data of Belgium pig herds and livestock movements. Federal Agency of the Safety of the Food Chain (FASFC), Belgium.

Sanson, R. L., Struthers, G., King, P., Weston, J.F. and Morris, R.S. (1993). The potential extent of transmission of foot-and-mouth disease: a study of the movement of animals and materials in Southland, New Zealand. N.Z. Vet. J. 41, 21-28.

Sanson, R.L. (2005). A survey to investigate movements off sheep and cattle farms in New Zealand, with reference to the potential transmission of foot-and-mouth disease. N. Z. Vet. J., 53, 223-233.

Stärk, K.D.C. (1998). Quantification of contacts between pig farms to assess the potential risk of classical swine fever transmission. In: Stärk, K.D.C., Systems for the prevention and control of infectious diseases in pigs. PhD thesis, Massey University, Palmerston North, NZ, pp. 202-215.

Stegeman, A., Elbers, A.R., Smak, J. and de Jong, M.C. (1999). Quantification of the transmission of classical swine fever virus between herds during the 1997-1998 epidemic in The Netherlands. Prev. Vet. Med. 42, 219-234.

Stegeman, A., Bouma, A., Elbers, A.R., de Jong, M.C., Nodelijk, G., de Klerk, F., Koch, G. and van Boven, M. (2004). Avian influenza A virus (H7N7) epidemic in The Netherlands in 2003: course of the epidemic and effectiveness of control measures. J. Infect. Dis. 190, 2088-2095.

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

AN ANALYSIS OF THE EFFECT OF THE INTRODUCTION OF PRE-MOVEMENT TESTING FOR BOVINE TB IN ENGLAND AND WALES

A. MITCHELL*, N. ALEXANDER, R. CLIFTON-HADLEY, J. COOPER, M. GILBERT, R. DE LA RUA-DOMENECH, J. MAWDSLEY AND W. WINT

SUMMARY

On 27 March 2006, the government introduced, with certain exemptions, the mandatory Pre-Movement Testing (PrMT) of cattle moving out of high bovine tuberculosis (bTB) risk herds in England (defined as those undergoing routine tuberculin testing every one or two years). Initially, cattle under 15 months of age were excluded. The same policy was introduced in Wales on 2 May 2006. Pre and post-movement testing had already been introduced in Scotland on 23 September 2005 for all cattle over 42 days old, moving into Scotland from parishes subject to annual or biennial tuberculin testing. On 1 March 2007, PrMT in England and Wales was further extended to animals over the age of 42 days.

This paper firstly describes a simple adaptation of the model developed by Gilbert et al. (2005) to predict the number of animals reacting to the tuberculin skin test and subsequently being confirmed to have bTB (confirmed reactors), that would be detected in the first year of PrMT. By reducing the number of cattle movements in the model it was possible to mimic and thus assess the effect of reducing the movement of infected cattle of different ages on the probability of confirmed bTB breakdowns, whilst still taking into account the influence of the other agricultural, environmental and climatic variables in the model. The model predictions provided have turned out to be remarkably accurate for phase 1 of PrMT. These predictions were then extended to cover the expected impact of extending PrMT to all cattle over 42 days old.

Secondly, the spatial distribution of bTB breakdowns first disclosed by PrMT was assessed. In the first year after the introduction of PrMT, the not unexpected finding, was that there was no clear relationship between areas where breakdowns were detected and the general changes in herd incidence of bTB in those areas, suggesting that PrMT was not yet having a marked effect on the national incidence of bTB.

The impact that PrMT has had on farmer behaviour thus far was also explored. From interrogation of the Cattle Tracing System (CTS) database, there is evidence that the effect of this new policy has been two-fold. Firstly, there has been a reduction in the number of cattle movements that would be eligible for PrMT. This is not surprising given that cattle owners bear the financial burden of PrMT. Secondly, there is also some evidence that cattle farmers have altered their farming practices to take advantage of the various exemptions to the PrMT regime,

* Andrew Mitchell, Centre for Epidemiology and Risk Analysis (CERA), Veterinary Laboratories Agency, New Haw, Addlestone, Surrey, KT15 3NB, UK. Email a.p.Mitchell@vla.defra.gsi.gov.uk

e.g. by moving cattle within the 60 days following a routine herd test paid for by the government.

Finally, the impact of further control measures on the spread of bTB by cattle movement was considered. For example, what the effect would be of: (i) extending PrMT to areas of the country which, by virtue of their current testing regime, are currently exempt; (ii) replacing or supplementing the tuberculin skin test presently used with the gamma-interferon (IFN) blood test and (iii) enhancing PrMT in England and Wales with post-movement testing.

INTRODUCTION

Bovine tuberculosis (bTB), caused by the bacterium *M. bovis*, is probably the most serious endemic disease of cattle in GB. Despite a steady increase in the number of infected herds reported year on year since the mid 1980s, the occurrence of the disease has remained geographically clustered, if steadily expanding. Only a few sporadic incidents (breakdowns) occur each year outside the core areas of high bTB incidence (the South West of England, the West Midlands and Wales).

Previous work such as that of Gilbert et al. (2005), Green et al. (2006) and Gopal et al. (2006), has shown that by far the most significant risk factor in the spread of bovine TB (bTB) from areas of high incidence to those of low incidence in Great Britain is the uncontrolled movement of potentially infected cattle. If the introduction of infected cattle into low incidence areas goes unchecked, it may eventually lead to the development of bTB ('hotspots') in areas previously considered free from disease. However, the precise circumstances that result in the establishment of such clusters in some areas and not in others are not fully understood. In addition to the persistence of a reservoir of *Mycobacterium bovi*s (*M. bovis)* infection in some badger populations of Great Britain (GB), it is believed that movements of infected cattle also contribute to the maintenance of bTB within areas of endemic TB incidence.

A major element of the Department for Environment Food and Rural Affairs (Defra) bTB strategy is to control the spread of the disease to new areas and to prevent the establishment of any new BTB 'hotspots' or endemic areas. Although there is a wildlife reservoir of infection in GB, it is the movement of infected cattle that is seen as posing a major risk for the spread of disease to these low incidence areas. Defra considered a number of policy options to lower the risk posed by cattle movement (Madders et al. 2005, Anon. 2004), including banning all cattle movements from high to low bTB incidence areas (zoning), and various combinations of pre- and post-movement tuberculin testing. The option adopted in England and Wales has been to test all animals before they moved out of an annually or biennially tested herd, regardless of the normal testing frequency of the destination herd. PrMT of cattle came officially into force in England and Wales on 27 March and 2 May 2006, respectively. However, in order to give some time to the cattle industry to adapt to the new policy, cattle under 15 months of age were initially exempted from PrMT, until 1 March 2007.

A number of movement categories have been exempted from PrMT since the outset, namely all cattle under 42 days old and those animals moving: (i) to slaughter (either directly or via certain types of approved markets or finishing units); (ii) in the 60 days following a clear tuberculin skin test; (iii) after they had been in a herd for less than 30 days; (iv) within a Sole Occupancy Authority (SOA); or (v) to an agricultural show, common grazing, semen collection centre or veterinary treatment centre.

After the first year of PrMT it was agreed that there should be a review of the policy: in particular there should be a cost-benefit analysis to determine whether PrMT should be extended to animals under the age of 15 months. As a result of this process, PrMT was duly extended to all animals over 42 days old in March 2007.

In GB, routine tuberculin testing of cattle herds takes place every one, two, three or four years according to the recent bTB history of the parish in which the herd resides, as determined by the Council Directive 64/432/EEC. The numbers of parishes, herds and animals covered by each testing regime is shown in Table 1. The PrMT regime was imposed on herds from high-risk parishes, defined as those undergoing one or two yearly bTB surveillance testing (mapped for 2007 in Figure 1). Within 3- and 4-yearly tested parishes, a small number of herds are subject to annual tuberculin testing (and thus PrMT) because they are deemed to pose significant animal health risk (e.g. bull hirers and cattle dealers). All cattle herds in Scotland are tested for bTB every 4 years.

Table 1. Number of Parishes, Herds and Animals in testing regimes in GB – November 2007

Inter-test Interval (years)		**1**	**2**	**3**	**4**	**TOTALS**
Parishes	**No.**	2,340	1,388	39	9,056	12,823
	%	18.25	10.82	0.30	70.63	100
Herds	**No.**	26,110	10,961	264	48,689	86,024
	%	30.35	12.74	0.31	56.60	100
Animals	**No.**	2,714,642	1,021,366	31,479	4,487,691	8,255,178
	%	32.88	12.37	0.38	54.37	100

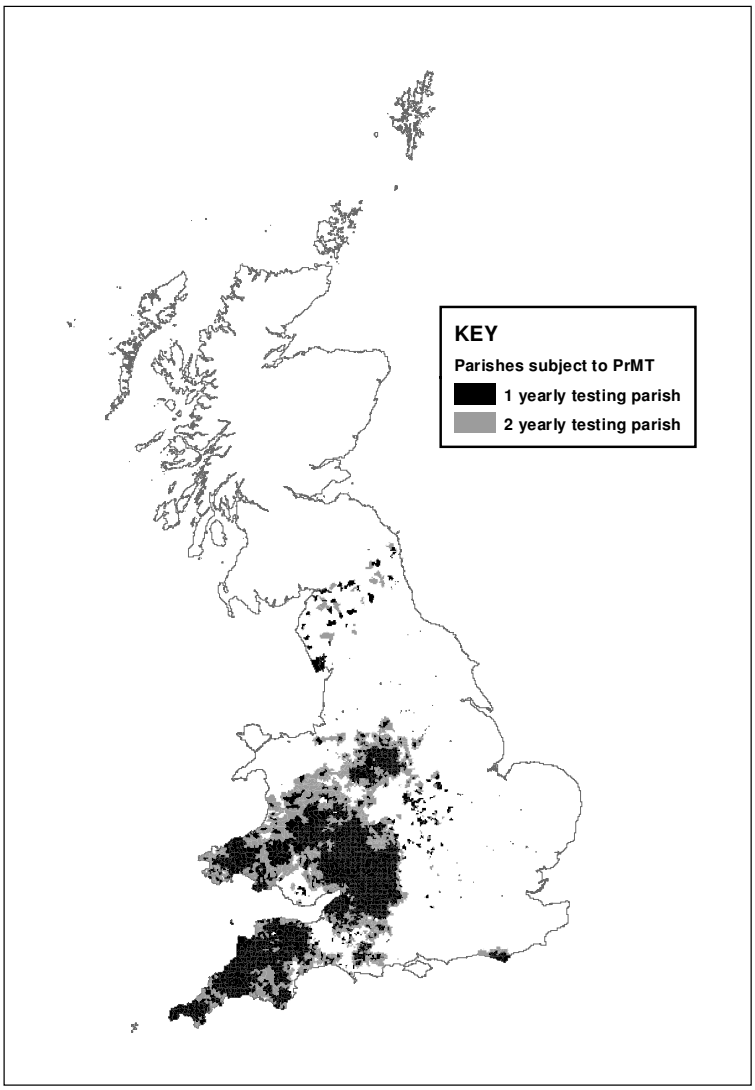


Fig. 1. Areas of GB subject to PrMT – November 2007

Bovine TB control policy is a devolved matter in GB. In order to protect the low bTB incidence status of Scotland, the Scottish Government adopted, from 23 September 2005, both pre- and post-movement testing for all cattle over 42 days old entering Scotland from 1 and 2-yearly testing parishes in England and Wales. In England and Wales only PrMT was implemented with the same criteria being applied in both countries, and coming into effect on 27 March 2006 in England and on the 2 May 2006 in Wales.

One significant difference between PrMT and other forms of mandatory bTB testing is that PrMT has to be arranged and paid for by the cattle owner, although tuberculin is supplied free of charge by the government. PrMT testing is performed (and certified) by the herd owner's own private veterinary surgeon or any other veterinarian approved by the government for the

administration of tuberculin tests. Therefore, PrMT charges vary between veterinary practices, although the unit cost of PrMT tends to fall sharply, as more animals are included in one PrMT event, because of the fixed costs associated with the test. The cost is also proportionally higher for young animals because of their lower stock value (Jinman et al., 2006). The precise impact of these costs on farmer behaviour was unknown when PrMT was introduced, although it was thought likely that some herd owners would modify their cattle movement patterns to minimise any additional costs incurred.

The purpose of this study was three-fold. Firstly, existing methods were used to model the likely effect of the first phase of PrMT on the number of bTB herd breakdowns and to assess the likely benefits of extending PrMT to other age groups of animals. Such work would allow Defra to ascertain whether PrMT was working as expected and the predictions provided for extending PrMT would allow them to undertake a cost-benefit analysis on extending the testing regime. A second objective was to assess the extent to which PrMT has actually contributed to overall bTB control. Thirdly, the effect that PrMT has had on farming practices was explored, as indicated by any changes to cattle movement patterns.

MATERIALS AND METHODS

To predict the consequences of PrMT, bTB distribution models developed by the authors (see Gilbert et al. 2005) were used. These were based on logistic regression analyses which established the statistical relationships between indicators of environmental, climatic and animal movement parameters and the distribution of the disease, defined as presence or absence of confirmed herd breakdowns within 5 by 5 kilometre grids. The models provide estimates of the probability that the disease (a new confirmed TB breakdown) will be present in each grid. An example for 2005 is shown in Figure 2.

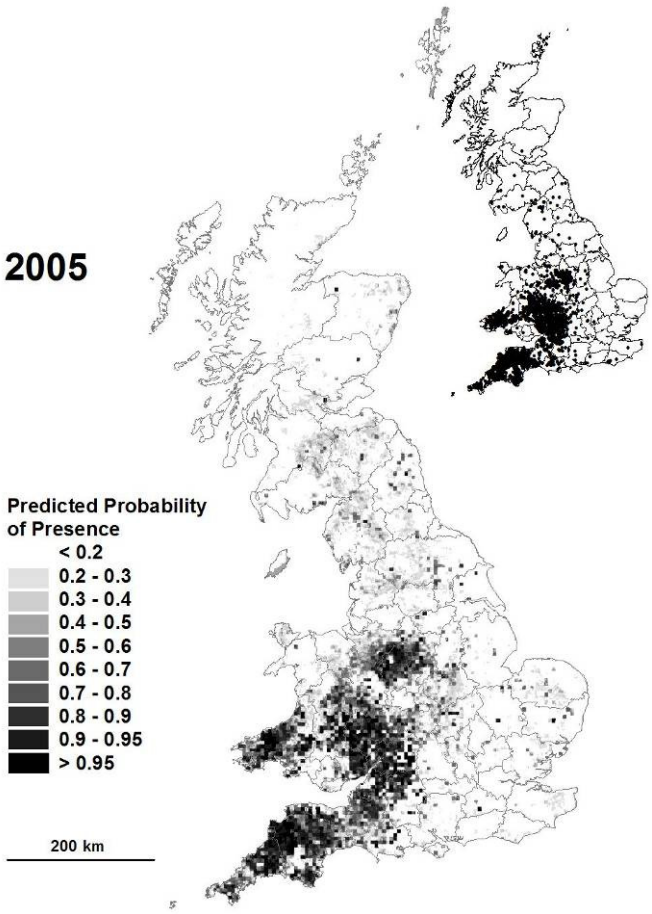


Fig. 2 Predicted model output - without PrMT (inset map represents the actual distribution of confirmed new breakdowns declared in GB throughout 2005

These models incorporate data from two extensive Defra databases, the first of which is the Cattle Tracing System (CTS), which contains records of births, deaths and movements for all cattle in Great Britain since 2000. The archive is of considerable size and complexity and the Veterinary Laboratories Agency (VLA) now has well-established methods for the processing and extraction of the data, which are described in Mitchell et al. (2005).

The bTB data, was obtained from VETNET - DEFRA's disease surveillance database, which records (amongst other things) details of bTB tests (including PrMT tests) and their results, as well as information on parish testing intervals. It should be noted, however that details of animals that test negative to PrMT are not recorded, and so the profile of such animals must be extracted from the CTS database, which can be interrogated to obtain the details of animals that should have been theoretically tested. However, this theoretical profile may not be exactly the same as that of the animals that are actually tested, due, for example, to non-compliance.

The original statistical models demonstrated that cattle movement patterns were the most significant predictor of disease distribution, particularly those cattle movements that originated from grids where confirmed bTB breakdowns had been identified (Gilbert et al. 2005). Initially, cattle movements were modelled regardless of animal age and origin location. To investigate the impact of PrMT, the models were modified to replace the original movement predictors with the movements of two age categories of cattle, moved from locations eligible for PrMT; i.e. those over 15 months old, and those between 42 days and 15 months from herds tested at 1 or 2 year

intervals. Movements of ineligible animals under 42 days old were also incorporated into the models. All of the predictions are based on bTB and cattle movement data for 2004, the most recent year for which a complete dataset was available when the models were first run in early 2006.

The model assumes that only 75% of infected animals would be detected as reactors by PrMT. This figure represents a conservative estimate for tuberculin test sensitivity at the individual animal level (de la Rua-Domenech et al. 2006). The model outputs were therefore calculated assuming the normal movement levels of each age group without PrMT and then re-run assuming an effective reduction of 75% in potentially infective movement numbers of the two cattle age groups affected by PrMT. Subtracting the second from the first provided a calculated percentage reduction in bTB probability per grid that could be ascribed to PrMT. The percentages for each grid were then multiplied by the recorded number of herd breakdowns per grid prior to PrMT, and these then summed for England and Wales to provide an estimate of the number of cases prevented.

The model, therefore, predicted the number of confirmed herd breakdowns that would not have occurred if 75% of the movements from 1 and 2 yearly tested grids where breakdowns had been previously recorded were stopped. The predicted impact thus assumes that movement patterns remain constant, and are not affected by the implementation of PrMT.

The only way of validating the model was to compare the predictions (in terms of breakdowns prevented), with the actual number of positive animals detected at PrMT. For the purposes of model validation it was assumed that detection of a confirmed reactor at PrMT would prevent one future bTB breakdown on a different holding. However, this assumption needs to be qualified. Although not all infected cattle are necessarily infectious at the time they are identified as reactors in a PrMT, there may be situations where one infectious animal causes more than one bTB breakdown e.g. by moving through (and infecting) several herds before it is detected by routine testing or meat inspection, or by infecting other animals in the destination herd which then move to another herd. The converse is also possible, whereby a batch of cattle containing more than one infectious animal are moved together and cause only a single bTB breakdown in the destination herd.

RESULTS

Model outputs for PrMT

The model predicts a reduction of 211 confirmed herd breakdowns in England and Wales in the first year of the PrMT policy, i.e. when all animals under 15 months were still exempt. The model predicted a further reduction of 246 confirmed breakdowns if animals between 42 days and 15 months old had been tested in that first year (see Table 2).

These predicted model outputs compare with an actual total of 221 confirmed reactors detected by PrMT during the first phase, i.e. when only animals over 15 months were tested. This represents a remarkably close match between predicted and actual figures, if it can be assumed that a single detection by PrMT would prevent, on average, a single breakdown.

Table 2: Predicted drop in bTB herd breakdowns resulting from the adoption of PrMT, data from 2004

	>15 months	42 days – 15 months	Total	Actual total confirmed breakdowns (2004)
England	170	208	378	1551
Wales	41	38	79	350
TOTAL	211	246	457	1901

Figures 3 and 4 provide the corresponding maps of grid estimates of percentage reduction on probability and calculated herd breakdown per grid prevented, from which these figures are aggregated. It is noticeable that the major impacts were predicted for the high bTB incidence areas rather than the three and four yearly tested areas. This may reflect the fact that most cattle movements within GB are comparatively short distance (Mitchell et al. 2006) and thus it is likely that most of the PrMT disclosures were in animals intended to move to nearby locations. Hence the number of new bTB breakdowns prevented would tend to be greatest in the high incidence areas of GB. This is considered further in the following sections.

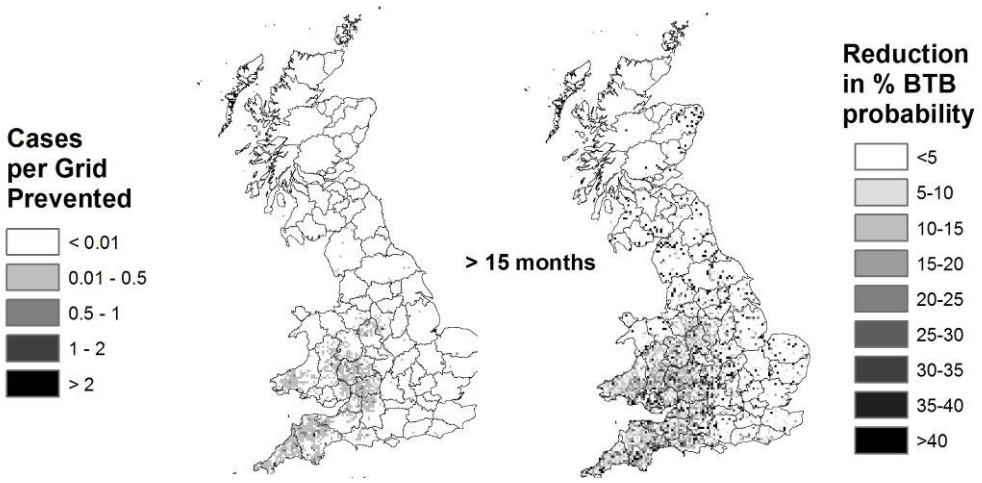


Fig. 3 Model outputs for animals > 15 months tested

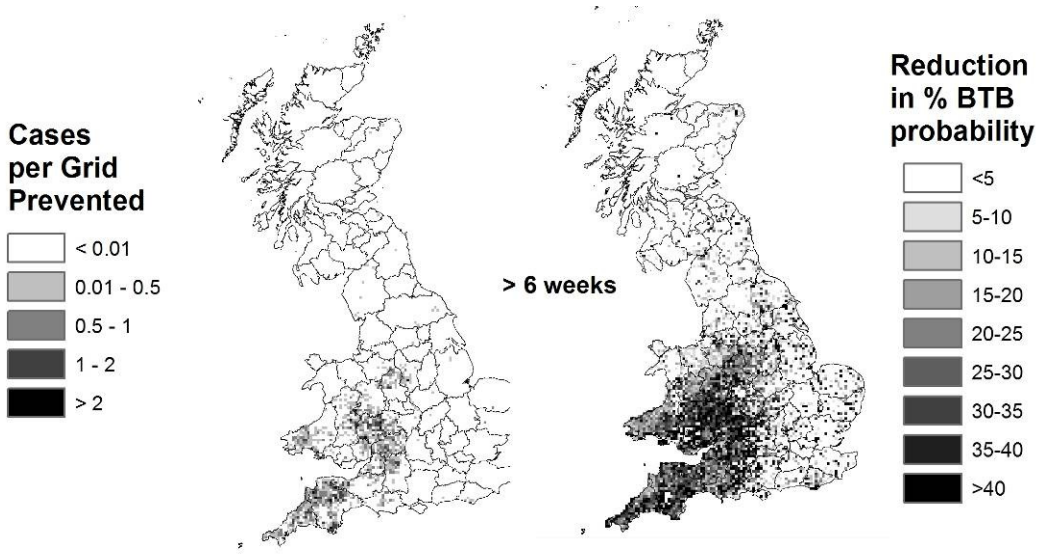


Fig. 4 Model outputs for all animals > 42 days tested

Figures 5 and 6 show that the gain from PrMT is proportional to the overall PrMT effort, with the numbers of both breakdowns and reactors detected at their highest in peak PrMT testing months. The 221 confirmed reactors detected by PrMT represented 56.8% of the 386 animals reacting to a PrMT during the first year of the policy. This compares favourably with a national average confirmation proportion approaching 40% for all tuberculin reactors, and suggests a higher than average positive predictive value for PrMTs, which by definition are carried out in areas of high bTB incidence.

Figures 7 and 8 both show the patterns of seasonality for PrMT disclosures and all other bTB testing in terms of both when the tests are being performed (Figure 7) and when such testing discloses breakdowns (Figure 8). Figure 7 shows that PrMT is a relatively small percentage of all testing carried out in any given month in 1- and 2-yearly testing areas. The peak months for bTB testing and breakdowns are in the autumn and winter, but this seasonal trend has not been mirrored for phase 1 of PrMT, with both PrMT testing and PrMT disclosed breakdowns showing a spring and autumn peak. Figure 8 shows that the distribution of all confirmed breakdowns in PrMT tested areas for 2005/06 and 2006/07 is relatively similar and has been unaffected by PrMT.

For Figures 5 - 8 : data relating to March refers only to 27/03/06 – 31/03/06. The period during which PrMT was in effect

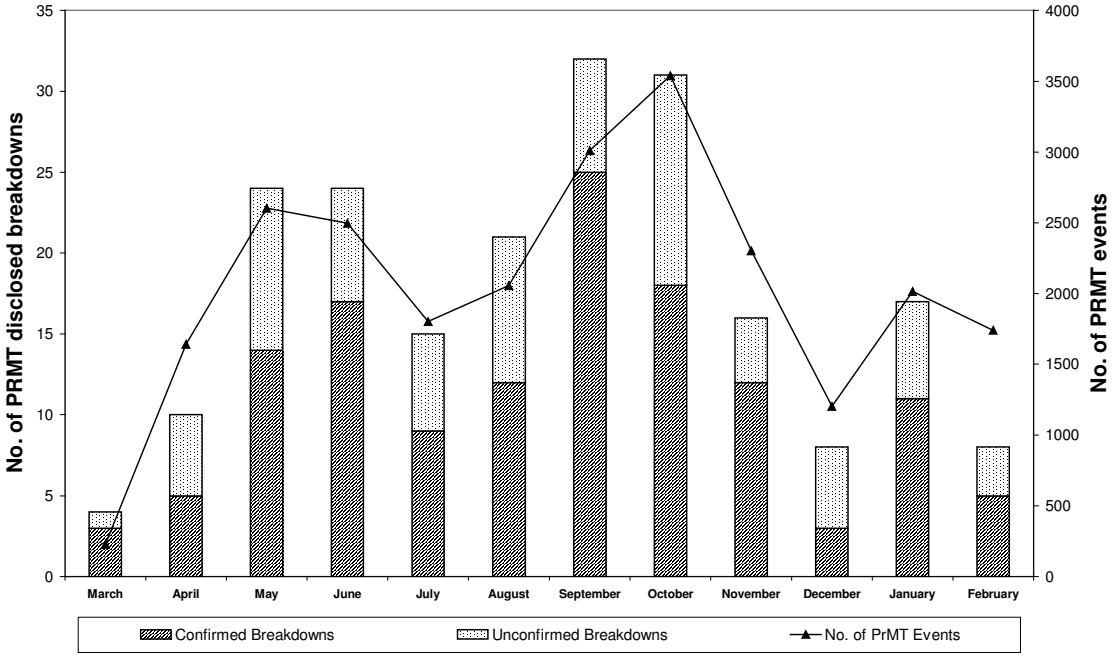


Fig. 5 Number of breakdowns detected by PrMT testing and number of herds tested during Phase 1 of PrMT

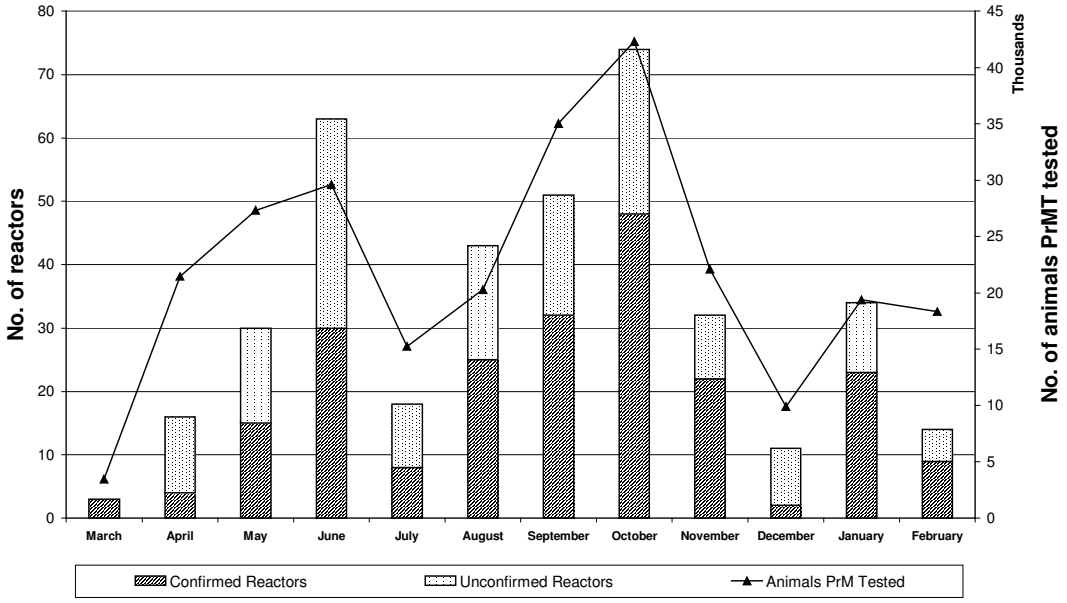


Fig. 6 Number of reactors detected at PrMT and number of animals tested

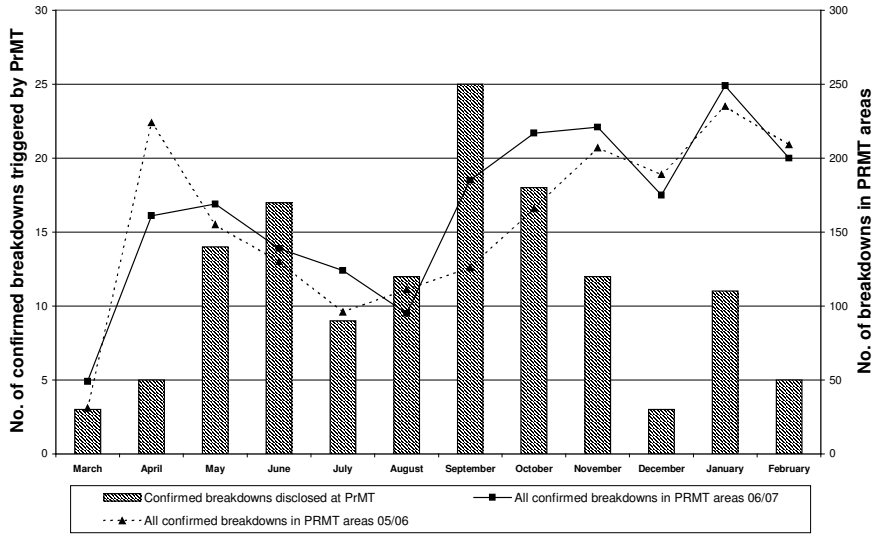


Fig. 7 Animals tested at PrMTs and in all other bTB tests in 1- and 2-yearly testing areas

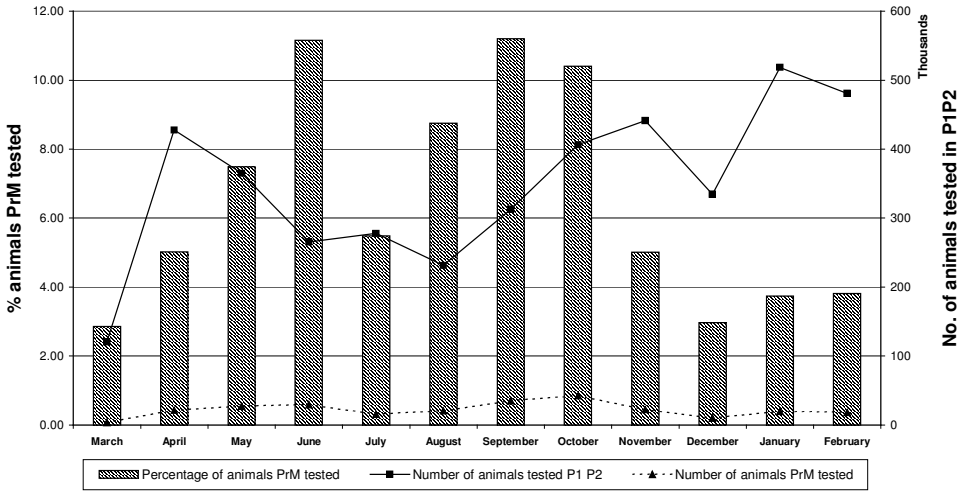


Fig. 8 Confirmed bTB breakdowns triggered by PrMT compared to all confirmed bTB breakdowns in 1- and 2-yearly testing areas.

Spatial distribution of PrMT testing and breakdowns

Although the primary aim of PrMT is to reduce the risk of translocating infected cattle between herds, a spin-off benefit of this policy is that it can enhance bTB surveillance in 1- and 2-yearly testing areas by disclosing infection in herds subjected to PrMT between two routine tests. Figure 9 shows the locations in which confirmed breakdowns were first disclosed by PrMT during the first year of implementation of the policy in England and Wales, with a backdrop of all confirmed breakdowns identified in the same period. It can be seen that there is no clear pattern to the confirmed breakdowns disclosed by PrMT, when compared to all other confirmed breakdowns in PrMT tested areas. However, when considered at a larger scale there were clearly some high-density breakdown areas with no breakdowns disclosed by PrMT.

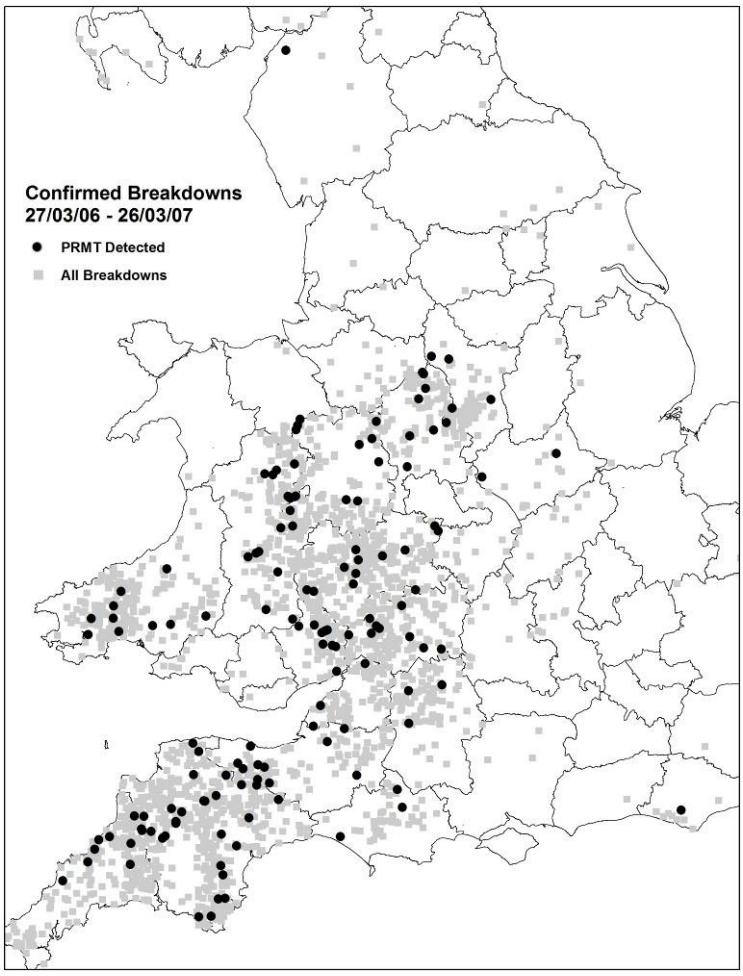


Fig. 9 Map of confirmed breakdowns triggered by a PrMT test

This spatial variation in breakdowns disclosed by PrMT is confirmed by Table 3, which shows that, at a county level, the effectiveness of PrMT in disclosing infection does vary. The percentage of confirmed breakdowns disclosed by PrMT in those counties listed ranges from 3.7 in Cornwall to 12.9 in Gwent, which suggests some real regional differences that are not just a reflection of the testing effort.

The effectiveness of PrMT, as an aid to stopping the spread of bTB to non-endemic areas, would be better understood if the intended destination of PrMT reactor animals was known. However this information is not recorded.

Table 3. County Distribution of PrMT disclosed breakdowns during phase 1

County**	PrMT Tests	Confirmed Breakdowns *	PrMT Breakdowns	% Breakdowns disclosed by PrMT	Breakdowns per 100 PrMT tests
Devon	3470	411	30	7.30	0.86
Hereford & Worcester	887	251	12	4.78	1.35
Dyfed	4230	192	11	5.73	0.26
Cornwall	1774	189	7	3.70	0.39
Powys	1981	152	15	9.87	0.76
Gloucestershire	672	148	9	6.08	1.34
Shropshire	1629	110	8	7.27	0.49
Staffordshire	1381	92	7	7.61	0.51
Avon	566	72	4	5.56	0.71
Wiltshire	718	70	4	5.71	0.56
Somerset	1201	66	7	10.61	0.58
Gwent	421	62	8	12.90	1.90
Derbyshire	1053	62	5	8.06	0.47

* Confirmed breakdowns in 1 or 2 yearly tested parishes in the counties listed.
** List limited to counties having 50 or more confirmed breakdowns during phase 1 of PrMT

Quantifying movements eligible for PrMT

The effectiveness of using actual PrMT disclosures to validate the predictive models, and to assess the number of breakdowns prevented by the testing regime, relies on the fact that cattle movement patterns have not been affected since the introduction of PrMT. If, however, the numbers of cattle moved between herds have been substantially reduced, then the impact of PrMT on bTB incidence might still be significant, even though the number of eligible movements (and thus the number of PrMT disclosures) would also be reduced.

Figure 10 shows movement figures from January 2004 until September 2007 obtained from the CTS database. For August and September 2007 the movement numbers were dramatically reduced by the impact of the Foot and Mouth Disease (FMD) outbreak in Surrey. Movements for each year are from areas defined as eligible for PrMT in 2006, to allow comparisons between years. Table 4 shows the overall numbers in each movement category with the percent change from the same period the previous year.

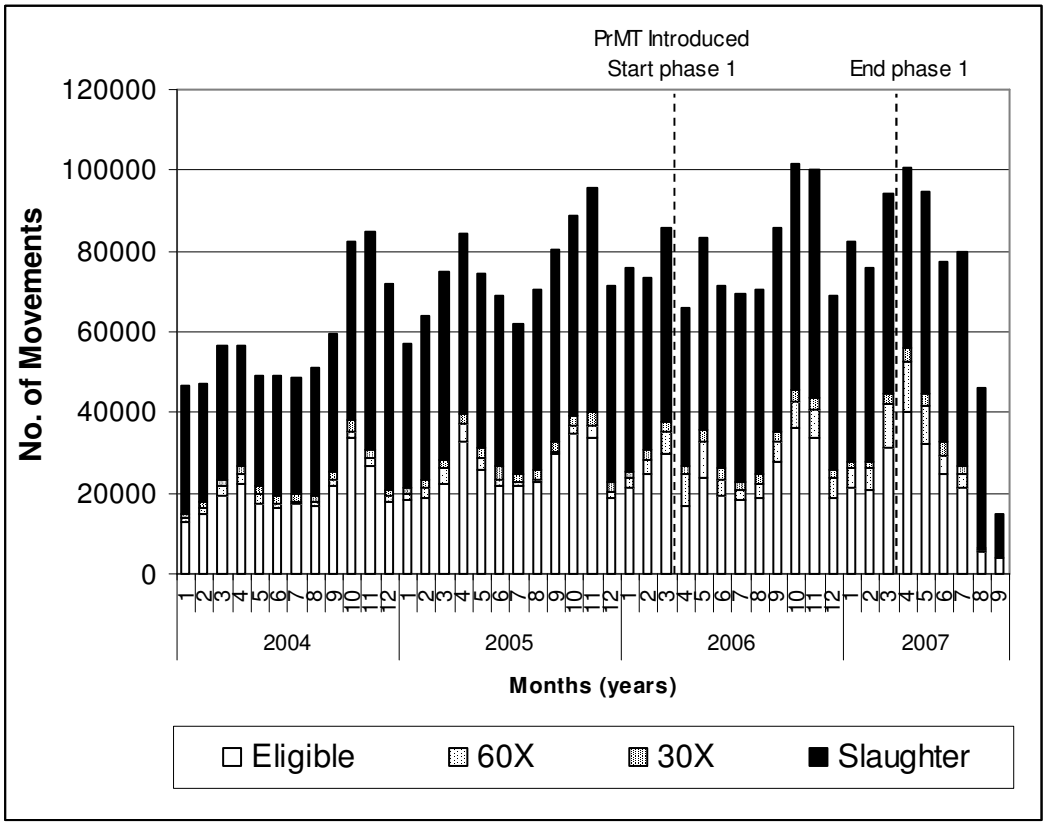


Fig. 10 Number of recorded cattle movements per month, for cattle aged over 15 months.

Table 4. Numbers of cattle over 15 months old moved during the first year of PrMT

Type of movement	Number (2006-07)	% of all movements	Number (2005-06)	% change on 05/06
Eligible	288187	29.70	317606	- 10
60 day exempt	70938	7.31	29134	+ 143
30 day exempt	27495	2.83	30070	- 9
To slaughter (exempt)	583586	60.15	554527	+ 5
TOTAL	970206	100.00	931337	+ 4

The data show that the introduction of PrMT whilst not having had an effect on overall movement numbers in PrMT areas, had an effect on cattle movement patterns. Herd owners are understandably looking to take advantage of the 60 day exemption when moving animals. It should also be noted that PrMT eligible movements account for a relatively small proportion of all movements, in the 15 month and older age group, because of the high percentage of animals

that are moving direct to slaughter. Such animals are, however, subject to routine carcass examination for bTB at the slaughterhouse, a less sensitive method of disclosing bTB infection than the tuberculin test (Corner et al., 1994)`.

The figure calculated from CTS as eligible for PrMT (288,187) is 8% more than those recorded on VetNet as actually Pre Movement tested (264,363). The additional movements most probably represent exemptions such as moves to slaughter markets and moves within Single Occupancy Authorities (SOAs) that were not identifiable from available data.

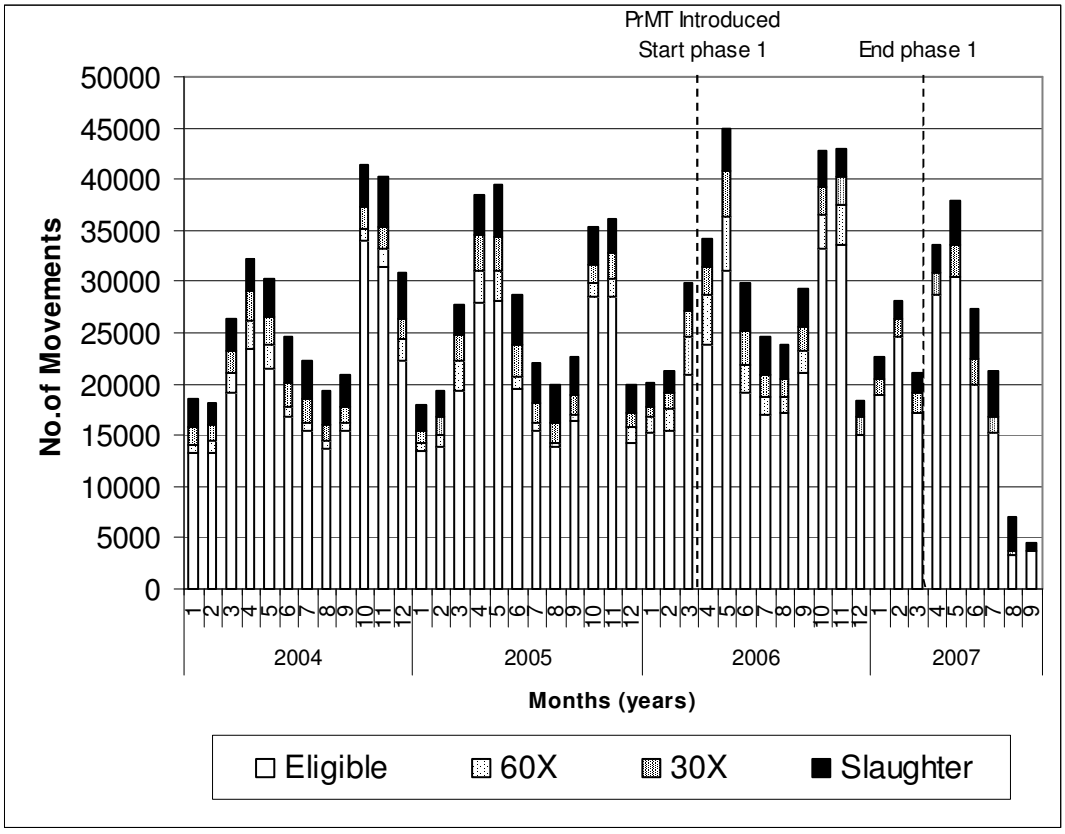


Fig. 11. Number of recorded cattle movements per month for animals between 42 days and 15 months old

Table 5. Numbers of cattle aged between 42 days and 15 months old moved during the first year of PrMT

Type of movement	Number (2006-07)	% of all movements	Number (2005-06)	% change on 05/06
Eligible	272057	72.37	244215	+ 11
60 day exempt	39343	10.47	21048	+ 87
30 day exempt	28938	7.70	26426	+ 9
To slaughter (exempt)	35577	9.46	42453	- 16
TOTAL	375915	100.00	334142	+ 12

Figure 11 and Table 5 show the movement figures for animals aged between 42 days and 15 months (42d-15m), to which PrMT was only extended in March 2007. Although the grand total of animals moving in this category is much less than for the >15 months category (375,915 as opposed to 970,206) because the percentage of animals moving direct to slaughter is much less, the number of "eligible" animals is broadly similar (272,057 as opposed to 288,187). Hence it can be expected that in the second year of PrMT, when the age restriction was lifted, the number of animals tested would be approximately doubled.

Table 5 shows that there has been no reduction in "eligible" animals during 2006 for animals in the 42d-15m category, which is to be expected because the policy was not yet in force. Interestingly there is also an increase in the number of animals in the 60 day exempt category (although not as large as in the > 15 month age range) this may be a result of animals in the 42d-15m category being moved at the same time as animals in the over 15 month category.

DISCUSSION

Although there is no way of directly validating the predictions provided at the start of phase 1 of PrMT, if the assumption is made that for every confirmed reactor detected (221) at PrMT one breakdown has been saved, then our prediction that 211 confirmed bTB breakdowns would be prevented has proved remarkably accurate.

The effect that PrMT has had on overall bTB levels is difficult to gauge. The last two years (2006 and 2007) have seen a stabilisation in overall bTB levels after steady rises in preceding years. However, PrMT is not the only bTB initiative introduced by Defra in recent years, other examples include: a year on year increase in surveillance testing; a stricter adherence to testing intervals (zero tolerance for overdue tests); and mandatory use of the gamma-interferon (IFN) blood test in certain circumstances. Given the relatively small numbers of animals tested during phase 1 of PrMT and the 2,083 new confirmed bTB breakdowns recorded in GB during 2005 (the year before PrMT introduction), it was thought unlikely to have had a dramatic impact on the overall bTB levels. Indeed, although the model predictions are in terms of "breakdowns saved", it is important to realise that in the first years of PrMT a reduction in breakdowns will not occur, the effect is simply to identify an infectious animal earlier – a breakdown still occurs

but on a different holding (the holding of origin, rather than on the holding of destination). However, if the model predictions for the full implementation of PrMT could be sustained over a number of years, then PrMT would undoubtedly result in reductions of bTB incidence on a national scale.

Much of the rationale behind the introduction of PrMT was to reduce the spread of bTB to non-endemic areas. It is difficult to assess whether PrMT is having the desired effect because of the absence of information on the intended destination of PrMT reactors, i.e. those animals detected may have been destined for low or high-risk areas. Even if such information were collected, approximately 50% of movements between farms in GB are via market (Mitchell et al. 2006), meaning the herd of destination would not have been known at the time of testing.

Analysis of bTB confirmed breakdowns, at a national scale, shows that in the time period covered by phase 1 of PrMT there were 113 confirmed breakdowns in 3 or 4 yearly parishes in England and Wales, this compares with 54 in the same period in 2005/06. It is of course impossible to say what the number of breakdowns in these low incidence areas would have been, without the implementation of PrMT. However, it would appear that PrMT has not yet had an impact on bTB incidence in low-incidence areas.

The spatial distribution of bTB breakdowns disclosed by PrMT show some variation at a county level that is not just a reflection of the testing effort within counties. The reasons for these variations are not clear, but may be a result of a number of factors: a reflection of the prevalence of bTB within the area; the different movement patterns within counties; changes in farmer behaviour following the introduction of PrMT and variable local levels of compliance with the legislation.

Analysis of CTS data shows that PrMT has led to changes in farming practice. In particular there is clear evidence that herd owners are seeking to take advantage of the 60 day exemption to move animals on the back of a prior herd tests, without incurring the cost of an additional PrMT. Whether the exemption is occurring by changes to the movement pattern or changes to the testing pattern is not yet known, although farmers are likely to have more freedom to change their movement patterns than their testing patterns. There is, however, no evidence that herd owners are looking to take advantage of the “30 day exemption”.

Analysis of CTS data shows that although the total number of movements for the 42-day – 15 month old category is considerably less than for those animals tested during phase 1, the number going direct to slaughter is much less, meaning the number of animals eligible for PrMT in the second year should be roughly double that in the first.

The models predict approximately 250 additional confirmed breakdowns would be prevented as a result of the extension of PrMT to all animals over the age of 42 days moving out of herds undergoing yearly and two-yearly bTB surveillance testing, leading to a total of 450 or roughly double the number predicted for the regime implemented in phase 1. As the incidence of bTB amongst younger animals is known to be less than that in older animals, this implies some critical and as yet unknown aspect of the movement of young stock in the spread of bTB, if the model predictions prove to be correct.

Although the model predicts a large number of additional breakdowns being prevented in phase 2, a large reduction in the movement of eligible stock in the 42d – 15m age group cannot be discounted. In phase 1 there was a 10% reduction in “eligible” movements and this reduction

may be greater for the 42d – 15m category, because of the proportionally higher cost of testing younger animals. This will reduce the apparent detection rate because fewer animals will be tested, although this should not affect the epidemiological impact because movement of infected animals will still be reduced.

During the second year of PrMT in England and Wales there have been two periods of restricted animal movement due to Foot and Mouth disease (FMD) and further movement restrictions imposed as a result of blue tongue. The consequence of these restrictions will be a dramatic reduction in the volume of cattle movements and consequently PrMT testing. This will undoubtedly dilute the direct impact of phase 2 of PrMT in the short term.

The choice of which areas should be subject to PrMT is obviously critical to its success. Defra's rationale behind the PrMT of animals moving from 1 and 2 yearly tested herds is that test interval is a proxy for bTB prevalence. In the main this is true, however the calculation of parish testing intervals (PTI) is reactive (i.e. the calculation to ascertain the PTI is based on historic data), so for new and emerging areas the PTI will not be representative of the infection in the area and consequently the area will not be subject to PrMT. Recent work by Cooper (in preparation) looking at the establishment of disease in East Carmarthenshire attributes several of the earliest breakdowns to movements out of 3 and 4 yearly parishes.

The recent final report of the ISG (Bourne et al. 2007), acknowledged that PrMT is a step on the road to addressing the problem of the spread of bTB through cattle movement. However, they recommended the use of the gamma-IFN test in parallel with the tuberculin test (to gain higher sensitivity) and in certain circumstances the use of post-movement testing, as initially put forward by the Madders et al. (2005). At present PrMT is based solely on the comparative intradermal tuberculin test, which this modelling work has assumed to have a sensitivity of 75%. A recent review of the diagnostic performance of the gamma-IFN test (de la Rua-Domenech et al. 2006) suggests that the median sensitivity of the IFN test is marginally greater than that of the comparative tuberculin test. Because the gamma-IFN and the skin test identify a slightly different population of *M. bovis*-infected cattle, the effect of using the tests in parallel would be to improve on the sensitivity of using either test alone, although as of yet, the combined sensitivity is unknown. It is expected that the models developed in this paper will provide valuable information for policy-makers on the predicted gain of using the tests in parallel.

The use of post-movement testing is already taking place for movements into Scotland and can also be privately requested in England. During the period of PrMT in England and Wales 6363 post-movement tests were carried out in GB with only one reactor (confirmed *M.bovis*) animal being disclosed. So, post-movement testing in addition to PrMT would not, in general, seem to represent a cost-effective policy option, although its use for the most high-risk cattle movements should not be discounted.

ACKNOWLEDGEMENTS

The authors would like to thank Defra for funding project SE3229 "Enhanced modelling and prediction of the spread of bovine TB in mainland Britain: impacts of cattle movements, climate and spoligotype" under which this work was done. The authors would also like to thank Defra, Professor Sir David Cox, Professor John Bourne and Barbara Wibberley for their helpful comments.

REFERENCES

Anon. (2004). Preparing for a new GB strategy on bovine tuberculosis - consultation document. Defra, London.

Bourne, J., Donnelly, C., Cox, D., Gettinby, G., McInerney, J., Morrison, I. and Woodroffe, R. (2007) Bovine TB: The Scientific Evidence. The Final report of the independent scientific group on cattle TB.

Cooper, J., Gilbert, M., Mitchell, A., Bourn, D., Mawdsley, J., Clifton-Hadley, R. & Wint, W. DEFRA project SE3229 Final report (in preparation)

Corner, L. A. (1994) Post mortem diagnosis of *Mycobacterium bovis* infection in cattle. Veterinary Microbiology. 40, 53-63.

de la Rua-Domenech R., Goodchild AV, Vordermeier HM, Hewinson RG, Christiansen KH, Clifton-Hadley RS. (2006). Ante mortem diagnosis of tuberculosis in cattle: a review of the tuberculin tests, γ-interferon assay and other ancillary diagnostic techniques. Research in Veterinary Science. 81, 190-210.

Gilbert, M., Mitchell, A., Bourn, D., Mawdsley, J., Clifton-Hadley, R. & 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., Clifton-Hadley, R., (2006). Introduction of bovine tuberculosis to north-east England by bought-in cattle. Veterinary Record, 159, 265-271.

Green, L.E and Cornell (2005). Investigations of cattle herd breakdowns with bovine tuberculosis in four counties of England and Wales using Vetnet data. Prev. Vet. Med., 70**,** 293-311.

Jinman, P., Madders, B., Jennings, B., Kirkwood, J. and Cunningham, A. (2006). Pre-Movement Testing. Report of the TB advisory group.

Madders, W., Attenborough, M., Beddoe, G., Brigstocke, T., Crea, H., Edwards, A., Messer-Bennetts, B., Nixon, R. and Padgett, C. (2005). DEFRA report - Introduction of a statutory requirement for bovine TB pre-movement testing in Great Britain.

Mitchell, A., Bourn, D., Mawsdley, J., Wint, W., Clifton-Hadley, R. and Gilbert, M. (2005) Characteristics of cattle movements in Britain – an analysis of records from the Cattle Tracing System. Animal Science. 80, 265-273.

DEVELOPMENT OF AN EARLY WARNING SYSTEM (EWS) FOR DETECTION OF CLASSICAL SWINE FEVER (CSF)

H. BROUWER-MIDDELESCH*, J.A. BACKER, W.A.J.M SWART, H.J.W VAN ROERMUND, VAN WOLF, AND G. VAN SCHAIK

SUMMARY

An Early Warning System (EWS) for Classical Swine Fever (CSF) based on mortality data was developed and the additional value of the EWS to the current surveillance programmes was determined. For this purpose daily pig mortality reports from 2001-2005, gathered by the Dutch rendering plant, were used. Three herd types were distinguished: finishing herds, sow herds and piglet herds. The mortality expressed as an incidence rate per animal-day at risk was highest in piglet herds (0.004 piglets/pig-day at risk), whereas the mortality incidence rate in finishing herds was only 0.0008 pigs/pig-day at risk. Sow herds had the lowest mortality incidence rate (0.0005 sows/pig-day at risk). To identify herds with an extremely high mortality, a number of absolute and relative thresholds in mortality for each herd type were determined. These thresholds were based on the mortality incidence rates. The thresholds were set at a higher level for small herds, because the mortality incidence rates were higher for small herds than for large herds. No real mortality data due to an outbreak of CSF were available to determine the sensitivity and specificity of the EWS. Therefore, simulated mortality data from a model (Backer et al., 2007a) were used. The additional value of the EWS was measured by determining the period of detection of CSF with the current surveillance programmes with and without incorporating the EWS based on mortality data (Backer et al., 2007b). The median time to detect a CSF infection in finishing herds and sow herds (assuming 50% mortality caused by CSFV) is two days earlier than with the current combination of surveillance programmes. The upper limit of the 95% confidence interval is also decreased by two days in finishing herds and by one day in sow herds. The EWS based on mortality data was not sensitive enough to detect the additional mortality due to CSFV in piglet herds (assuming 100% mortality caused by CSFV), probably due to already high and variable mortality rates in piglets under normal circumstances. In addition, piglet mortality and herd size are uncertain factors and therefore mortality incidence rates are less reliable for piglet herds. The effectiveness of the EWS based on mortality data to detect CSF outbreaks might increase when mortality data are complete and when the alertness of the farmer and veterinarian decreases. In addition, mortality data of the Dutch pig industry are objectively gathered and not influenced by the willingness of a farmer to report high pig mortality. The EWS could also have value for the detection of other diseases, causing mortality.

INTRODUCTION

*H. Brouwer-Middelesch Department of Diagnostics, Research & Epidemiology, GD Animal Health Service Ltd., Deventer tel: +31610960236; fax: +31570634104, E-mail address: h.middelesch@gddeventer.com

Classical swine fever (CSF) is a serious viral animal disease affecting pigs and wild boar. The entry of classical swine fever virus (CSFV) into non-vaccinated pig populations can cause major outbreaks with high economic losses and severe consequences for animal trade and animal welfare. The domestic pig population of the European Union (EU) is not preventively vaccinated against CSFV, because importing countries do not accept vaccinated pigs. Early detection and response to a suspected CSF outbreak is of major importance to maximize the effectiveness of control measures taken and to minimize the decrease in animal welfare and animal trade and the costs associated with an outbreak (Murray and McCutcheon, 1999; Mangen and Burrell, 2003). The EU non-vaccination policy is therefore based on stamping out and intensive surveillance.

In The Netherlands several surveillance programmes exist for the control and early detection of CSF: 1) daily clinical observation by the farmer; 2) monthly clinical inspection by a veterinarian; 3) testing blood samples by polymerase chain reaction (PCR) for CSF of diseased pigs; 4) routine pathology of severely diseased pigs and 5) routine virological tests on tonsils of all severely diseased pigs submitted for pathology. These surveillance programmes are all initiated by the onset of clinical symptoms of CSF. However, in general CSF symptoms are non-specific, especially at the early onset of the disease and therefore may be misinterpreted as symptoms of endemic diseases (Moennig, 1990; Elbers et al., 2002). Late detection of the presence of a CSF infection is mainly due to non-specific clinical symptoms (Dewulf et al., 2001). Misinterpretation of CSF symptoms extends the High Risk Period (HRP) that is the time between CSFV introduction in the country and the first detection of an infected farm, with CSFV having more time to infect other farms and consequently higher costs are needed to control the outbreak.

The aim of this study was to develop an Early Warning System (EWS) for the detection of CSF in The Netherlands based on mortality data gathered by the national rendering plant. The advantage of such a system is that it is independent from the cooperation of pig farmers and veterinarians and all herds would participate automatically. In addition, reporting animal cadavers to the rendering plant is obligatory in The Netherlands for all animals except for pets and horses. The additional value of the EWS to the current surveillance programmes is determined in this study with the use of simulated data of a CSF outbreak by Backer et al. (2007a,b) by determining the period of detection of CSF with and without incorporating the EWS based on mortality data.

MATERIAL AND METHODS

Description of Rendac data

The Dutch rendering plant “Rendac” is the only institution in The Netherlands that is allowed to collect and destroy animal cadavers. Rendac collects animal cadavers during week-days and normally no animal cadavers are collected during the week-ends. The herd owner reports animal cadavers to Rendac and when reported before 03.30 am, Rendac will collect the cadavers the same week-day.

To develop an EWS for CSF based on mortality data, all daily pig mortality reports over the period January 2001-December 2005 (a CSF-free period) registered by Rendac were used. This dataset included a unique Rendac client number of the herd owner reporting the cadaver, which is in most cases the same as unique herd identity number (UHI), day of cadaver collection by

Rendac, cadaver type (sow/boar, finishing pig or piglet) and the number of cadavers reported by the herd owner. In addition, data from the monthly clinical inspections over the period January 2001-December 2005 including UHI and herd size (number of sows/boars and finishing pigs) were used. Before linking herd size to the daily pig mortality reports, mean herd size per quarter was determined, because not every UHI had a clinical inspection every month. Both datasets were linked by UHI and time-unit i.e. all pig mortality reports from one UHI within a quarter got the same herd size. Herds with a mean herd size of less than 10 pigs were not included in the analyses, because of possible administrative errors or discrepancy between Rendac client number and UHI. As a result, 546 finishing herds (3.7%), 529 piglet herds (7.2%) and 1,434 sow herds (17.2%) were excluded from the analyses.

In a descriptive analysis of the Rendac data the following key monitoring indicators were determined over the period January 2001-December 2005:

- Percentage of herds reporting dead pigs to Rendac per week-day;
- Number of dead pigs per report; and
- Mortality incidence rate

Mortality incidence rates (expressed as the number of dead pigs per pig-day at risk) per report were determined by dividing the number of daily reported dead pigs for each herd by the number of days between consecutive reports and the number of pigs present in a herd:

$$\text{Mortality incidence rate}_{i,j} = \frac{m_{i,j}}{(day_{i,j} - day_{i,j-1}) * \frac{(N_{i,j} + N_{i,j-1})}{2}} \quad (1)$$

where m is the number of dead pigs reported by the herd owner to Rendac, N is the number of pigs present in a herd and i is the herd owner with report j.

Differences in mortality incidence rates between herd types (i.e. finishing herds, sow herds and piglet herds) were analysed using the Kruskal Wallis test. An estimation was made for the number of piglets present in a herd, because the number of piglets were not registered during the monthly clinical inspections. Therefore, the number of piglets present in a herd was determined by Backer et al. (2007a,b) assuming a sow has 2.4 farrows a year and a mean litter size of 11.5 piglets, of which 10.5 piglets will survive in the first days. The estimated time present on the farm for piglets was assumed to be 63 days. Using these assumptions, a ratio of 4.3 piglets per sow (2.4 farrows*10.5 piglets per farrow*63 days present/365 days) was estimated. This proportion was comparable to the proportion of 4.4 found by Klinkenberg et al. (2003). In addition, no distinction was made between sows and boars during the clinical inspections. However, the number of boars present in a herd is usually very small in Dutch pig herds and thus the herd size was assumed to consist of sows only.

Differences in mortality incidence rates between small and large herds were analysed using the Wilcoxon's test. Cut points for the size of small and large herds were chosen in such a way that mortality incidence rates were not influenced by herd size above these cut points (Table 1).

Table 1. Number of pigs present in small and large herds per herd type in The Netherlands

Herd type	Small herds	Large herds
Finishing pigs	<200	≥200
Piglets	<400	≥400
Sows	<200	≥200

For every herd type (i.e. finishing, sow and piglet herds) and herd size (i.e. small and large) the percentage of herds reporting dead pigs to Rendac per week-day were determined. Differences in the number of dead pigs per report between herd types were analysed using the Kruskal Wallis test, whereas differences in the number of dead pigs per report between small and large herds were analysed using the Wilcoxon's test.

The mean number of herds with an extremely high pig mortality was determined every report-day, using a number of absolute and relative thresholds in mortality incidence rates. All data analyses were performed using SAS® 9.1.2, (SAS Institute Inc., Cary, NC, USA).

Additional value of EWS based on mortality incidence rates to the current surveillance programmes

No real-life CSF mortality data were available to determine the sensitivity and specificity of the EWS based on mortality incidence rates corrected for herd size. Therefore, simulated data from a model by Backer et al. (2007a) were used. The transmission of CSFV between pigs within a pen was described with a SEIR model, where S is the number of susceptible, E is the number of latently infected, I is the number of infectious and R is the number of removed (recovered or dead) pigs. In addition, the development of disease symptoms and the mortality due to a CSF outbreak is simulated. The additional value of the EWS was measured by determining the time to detection of CSF with the current surveillance programmes (detection by veterinarian and herd owner) with and without incorporating the EWS based on mortality data (Backer et al., 2007b).

For the EWS, daily pig mortality reports from the Dutch pig industry gathered by the rendering plant over a period from January 2001-December 2005 were used to fit the model for a situation without CSF. Subsequently, an outbreak of CSF was simulated and the value of the mortality thresholds, based on a situation without CSF, was determined. The analysis distinguished between herd type and for every herd type small and large herds were distinguished.

Based on experimental studies (Laevens et al., 1998; Moormann et al., 2000; Bouma et al., 2000) a mortality of 100% due to CSF was assumed in piglet herds. The mortality due to CSF was 40% in finishing herds and 20% in sow herds using a mild virulent isolate "souche Lorraine" (Dewulf et al., 2001; Dewulf et al., 2001b). High virulence isolates will probably cause higher CSF mortality rates and therefore the mortality in finishing herds and sow herds was set at a level of 50%.

Sensitivity and additional value of an EWS based on mortality rates without a correction for herd size

Since the 1st of January 2006 the numbers of pigs present in a herd, determined during the monthly clinical inspections, are no longer centrally gathered and herd size is unknown since that time. Therefore, the EWS for CSF was also based on mortality rates *without* a correction for herd size thus, the number of daily reported dead pigs for each herd was divided by the number of days between consecutive reports (expressed as the number of dead pigs per day):

$$\text{Mortality rate}_{i,j} = \frac{m_{i,j}}{(day_{i,j} - day_{i,j-1})} \quad (2)$$

where m is the number of dead pigs reported by the herd owner to Rendac and i is the herd owner with report j.

The sensitivity of the EWS based on mortality rates *without* a correction for herd size was determined. The numbers of herds with normal and extremely high pig mortality over the period January 2001-December 2005 were compared using mortality thresholds based on mortality incidence rates *with* a correction for herd size and mortality rates *without* a correction for herd size. The EWS based on mortality incidence rates *with* a correction for herd size was taken as gold standard. The additional value was also determined in the model of Backer et al. (2007b), but those results are not shown in this paper.

RESULTS

Descriptive analysis of Rendac data

Percentage of herds reporting dead pigs to Rendac

In the period 2001-2005, the mean percentage of herds reporting dead pigs to Rendac was lower in finishing than in piglet and sow herds. The mean percentage of herds reporting dead pigs to Rendac was the same for piglet herds and sow herds. More large herds reported dead pigs to Rendac every weekday than small herds (Table 2).

Table 2. Descriptive analysis of the mean number of herds and the percentage of herds reporting dead pigs to Rendac every week-day by herd type and size, 2001-2005

Herd type	Mean number of herds	Percentage of herds reporting dead pigs to Rendac per week-day
Finishing pigs		
All herds	10,649	8.4%
Small herds (<200)	3,668	3.8%
Large herds (≥200)	6,981	10.7%
Piglets		
All herds	4,946	10.3%
Small herds (<400)	1,282	4.8%
Large herds (≥400)	3,664	12.2%
Sows		
All herds	5,116	10.2%
Small herds (<200)	2,671	7.8%
Large herds (≥200)	2,445	12.8%

For all herd types the percentage of herds reporting dead pigs (i.e. finishing pigs, piglets or sows) to Rendac were constant in time. In the months June till August more herds reported dead pigs to Rendac (Fig. 1: only data from herds reporting dead piglets are shown).

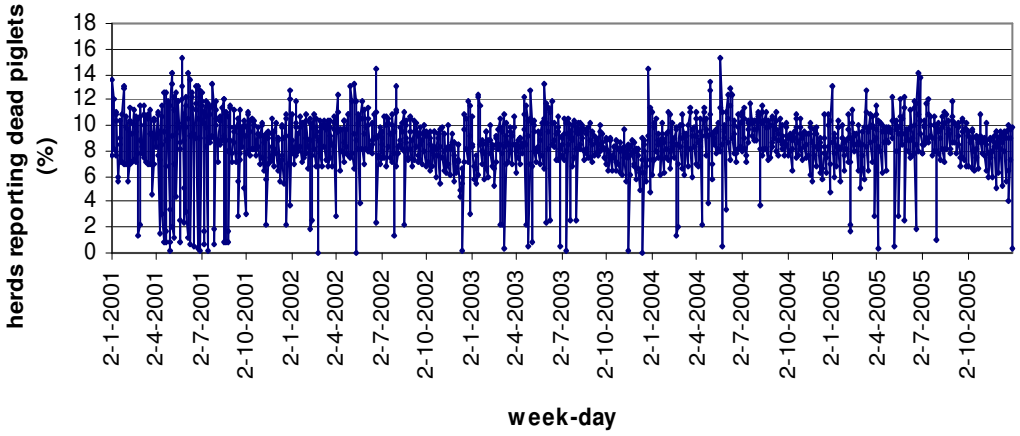


Fig. 1. Percentage of Dutch herds reporting dead piglets to Rendac every week-day, 2001-2005

Number of dead pigs per report

The number of dead pigs per report was significantly different between herd types (P<0.0001) (Table 3). Small herds reported significantly fewer dead pigs per report than large herds (P<0.0001).

Table 3. Descriptive analysis of the number of dead pigs per report by herd type and size in The Netherlands, 2001-2005

Herd type	Number dead pigs per report				
	Mean	Median	IQR*	99-percentile	Maximum
Finishing pigs					
All herds	2.7	2.0	2.0	12	1,771
Small herds (<200)	1.7	1.0	1.0	7	220
Large herds (≥200)	2.9	2.0	3.0	12	1,771
Piglets					
All herds	44.4	35.0	35.0	200	8,800
Small herds (<400)	15.2	10.0	15.0	78	525
Large herds (≥400)	45.4	35.0	34.0	200	8,800
Sows					
All herds	1.3	1.0	0.0	4	210
Small herds (<200)	1.1	1.0	0.0	3	210
Large herds (≥200)	1.3	1.0	0.0	4	165

*inter-quartile range

The mean number of dead piglets per report increased in time from 38.5 dead piglets per report in 2001 to 48.9 dead piglets per report in 2005. Herds with finishing pigs and sows reported most dead pigs in 2002 (3.0 dead finishing pigs versus 1.4 dead sows per report). The mean number of dead sows per report remained steady after 2002 with 1.2 dead sows per report, whereas the mean number of dead finishing pigs decreased in time from 2.8 dead finishing pigs per report in 2003 to 2.6 dead finishing pigs per report in 2005 (Fig. 2).

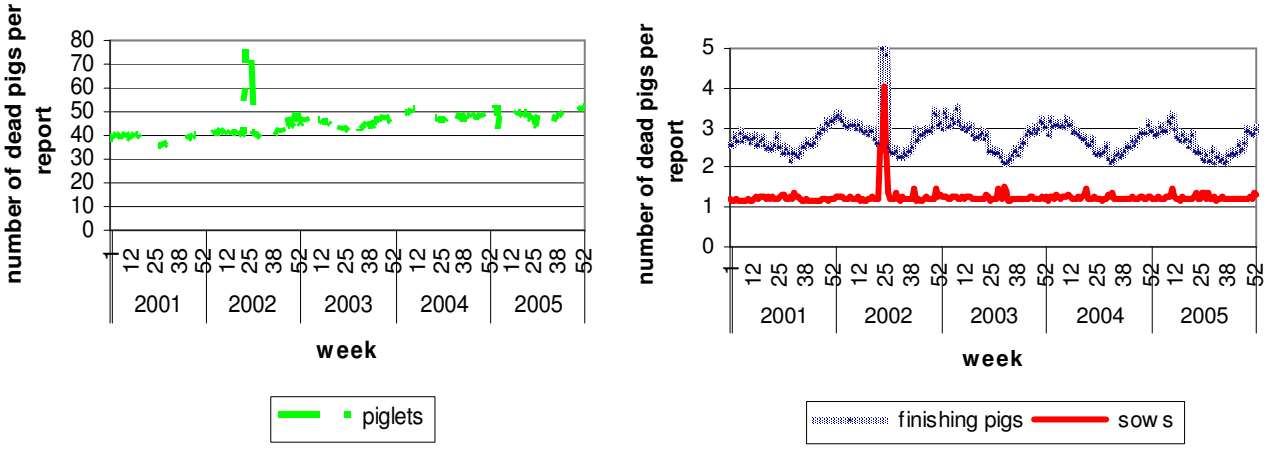


Fig. 2. Mean number of dead pigs per report by herd type in The Netherlands, 2001-2005

Mortality incidence rates per report

Mortality incidence rates were significantly different between herd types (P<0.0001) (Table 4). In addition, small herds had significantly higher mortality incidence rates than large herds (P<0.0001).

Table 4. Descriptive analysis of mortality incidence rates (no. dead pigs/pig-day at risk) by herd type and size in The Netherland, 2001-2005

Herd type	Mortality incidence rate				
	Mean	Median	IQR*	99-percentile	Maximum
Finishing pigs					
All herds	0.0008	0.0004	0.0005	0.0007	1.03
Small herds (<200)	0.0028	0.0008	0.0020	0.0330	1.03
Large herds (≥200)	0.0006	0.0004	0.0005	0.0036	0.82
Piglets					
All herds	0.0041	0.0032	0.0026	0.0178	1.27
Small herds (<400)	0.0079	0.0042	0.0056	0.0668	1.27
Large herds (≥400)	0.0040	0.0032	0.0026	0.0158	0.42
Sows					
All herds	0.0005	0.0002	0.0004	0.0041	1.91
Small herds (<200)	0.0007	0.0003	0.0005	0.0063	1.91
Large herds (≥200)	0.0004	0.0002	0.0003	0.0032	0.088

*inter quartile range

For all herd types the mean mortality incidence rate decreased in time. Mortality incidence rates within finishing herds decreased from 0.0010 dead pigs/pig-day at risk in 2001 to 0.0006 dead pigs/pig-day at risk in 2005. In the period 2001-2005 mortality incidence rates within piglet herds decreased with 0.0006 dead piglets/pig-day at risk (from 0.0044 dead piglets/pig-

day at risk in 2001 to 0.0038 dead piglets/pig-day at risk in 2005). In addition, mortality incidence rates within sow herds decreased from 0.0005 dead sows/pig-day at risk in 2001 to 0.0004 dead sows/pig-day at risk in 2005. Remarkably, for all herd types a peak in the mortality incidence rate was seen during the summer of 2002 (Fig. 1). This peak was caused by the medroxyprogesteronacetate (MPA) affair, where MPA contaminated pig feed was fed in a number of herds. Sows present in these herds could not farrow and were culled, whereas finishing pigs were preventively culled to protect public health.

A peak in mortality incidence rate was also seen in sow herds during the summer of 2003, which was probably caused by heat stress. This peak was not seen in finishing herds or piglet herds, because finishers and piglets are less vulnerable to high temperatures than sows (Fig. 3).

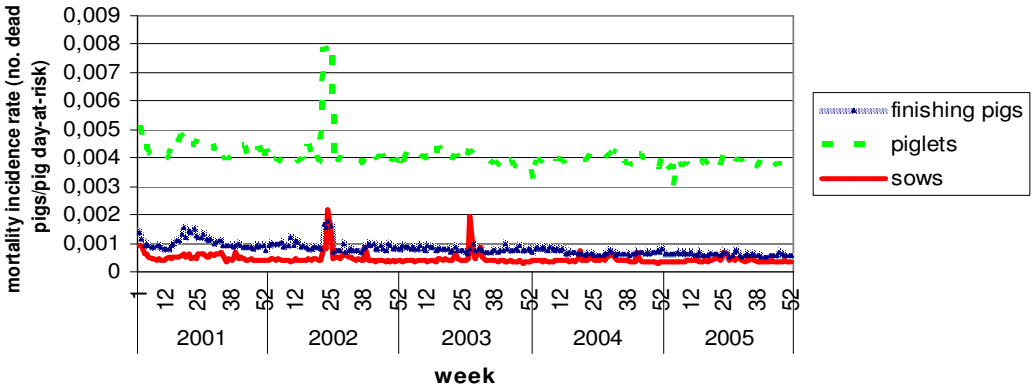


Fig. 3. Mean mortality incidence rates by herd type in The Netherlands, 2001-2005

EWS for CSF based on mortality data

The EWS for CSF using mortality data was based on mortality incidence rates per report. A distinction was made between finishing herds, piglet herds and sow herds, because mortality incidence rates and report patterns were different between those herd types. In addition, a distinction was made between small and large herds, because mortality incidence rates were higher in small herds than in large herds.

To identify herds with an extremely high pig mortality every report-day, a number of relative and absolute thresholds in mortality incidence rates were determined for each herd type and herd size (i.e. small and large herds) based on mortality data over the period 2001-2005. These thresholds were based on the mortality incidence rate on the day of the report (D) as well as on the differences between mortality incidence rates on report-day D and report-day D-1, report-day D-2 and report-day D-4.

Relative thresholds were determined every report-day. A number of relative thresholds were used and compared:

- The value of the 0.5% herds with the highest mortality incidence rates and/or difference between mortality incidence rates per report-day (top 0.5%);
- The value of the 1% herds with the highest mortality incidence rates and/or difference between mortality incidence rates per report-day (top 1%); and
- The value of the 2% herds with the highest mortality incidence rates and/or difference between mortality incidence rates per report-day (top 2%).

The absolute thresholds, shown in Table 5, were based on the value of 0.1% herds with the highest mortality incidence rates per herd type and herd size over the period 2001-2005.

Table 5. Absolute thresholds in mortality incidence rates by herd type and size in The Netherlands, 2001-2005

	Small herds	Large herds
	Finishing pigs (<200)	Finishing pigs (≥200)
Mortality incidence rate $_{\text{day D}}$	0.107	0.010
Difference between mortality incidence rate on day D and day D-1	0.089	0.008
Difference between mortality incidence rate on day D and day D-2	0.077	0.008
Difference between mortality incidence rate on day D and day D-4	0.082	0.008
	Piglets (<400)	Piglets (≥400)
Mortality incidence rate $_{\text{day D}}$	0.097	0.040
Difference between mortality incidence rate on day D and day D-1	0.067	0.034
Difference between mortality incidence rate on day D and day D-2	0.073	0.033
Difference between mortality incidence rate on day D and day D-4	0.085	0.033
	Sows (< 200)	Sows (≥ 200)
Mortality incidence rate$_{\text{day day D}}$	0.017	0.008
Difference between mortality incidence rate on day D and day D-1	0.016	0.008
Difference between mortality incidence rate on day D and day D-2	0.016	0.008
Difference between mortality incidence rate on day D and day D-4	0.017	0.008

When the mortality incidence rate of a report or difference in mortality incidence rates between reports was above at least one of the four relative or four absolute thresholds, the herd reporting the dead pigs would be identified as a herd with an extremely high pig mortality.

Mean number of herds identified with an extremely high mortality

Table 6 shows that herds reporting dead finishing pigs were more often identified as herds with an extremely high pig mortality than herds reporting dead piglets or sows. In addition, more large herds were identified as herds with an extremely high pig mortality than small herds. When relative thresholds were doubled (from 1% to 2%) or halved (from 1% to 0.5%), the number of herds with an extremely high pig mortality was also doubled or halved.

Table 6. Number of herds identified with an extremely high pig mortality per report day by herd type and size in The Netherlands, 2001-2005

	Absolute+ top 0.5%		Absolute + top 1%		Absolute + top 2%	
	mean	maximum	mean	maximum	mean	maximum
Finishing pigs						
Small herds (<200)	1.8	4	2.4	7	4.0	10
Large herds (≥200)	6.2	13	11.5	26	22.2	47
Piglets						
Small herds (<400)	1.9	4	1.9	4	1.9	4
Large herds (≥400)	4.1	10	7.6	18	14.7	28
Sows						
Small herds (<200)	1.6	10	1.6	10	2.0	10
Large herds (≥200)	1.7	37	2.7	37	4.2	37
Total	17.3		27.7		49	

Additional value of EWS based on mortality incidence rates to the current surveillance programmes

Table 7 shows that the median time to detect a CSF outbreak in finishing herds with the current surveillance programmes. Including the EWS based on mortality incidence rates, the median time to detect a CSF outbreak and the 2.5% most extreme detection times (97.5% percentile) would be shortened by two days (assuming 50% mortality caused by CSFV) when using relative thresholds based on the value of the 1% herds with the highest mortality incidence rates and/or difference between mortality incidence rates. The 97.5% percentile decreased by only two days when the number of finishing herds with an extremely high pig mortality was set at 1% or 2% and by one day when the number of finishing herds with an extremely high pig mortality was set at 0.5%.

The median time to detect a CSF outbreak in piglet herds was 32 days with the current surveillance programmes and was unchanged after including the EWS based on mortality incidence rates (assuming 100% mortality caused by CSFV). The 97.5% percentile decreased by one day when relative thresholds for the number of piglet herds with an extremely high pig mortality were set at 1% or 2% (Table 7).

The median time to detect a CSF outbreak in sow herds was 42 days with the current surveillance programmes. Including the EWS based on mortality incidence data, the median time to detection of CSF would be shortened by two days (assuming 50% mortality caused by CSFV) when using relative thresholds based on the value of the 1% herds with the highest mortality incidence rates and/or difference between mortality incidence rates. However, the 2.5% most extreme detection times decreased by only one day (Table 7). When the EWS based on mortality incidence rates was included, a CSF outbreak would be detected earlier in large herds than in small herds (Table 7).

Table 7. Time between introduction of CSFV in the country and first detection of a CSF infected herd (median with 95% CI) by herd type and size for the current surveillance programmes with and without the EWS based on mortality incidence rates.

	Current surveillance programmes	Current surveillance programmes + Early Warning System based on mortality data		
		top 0.5%	top 1%	top 2%
Finishing pigs				
All herds	38 (27 – 53)	37 (25-52)	36 (24-51)	34 (23-51)
Small herds (<200)	39 (27 – 56)	39 (26-56)	37 (25-55)	36 (24-54)
Large herds (≥200)	38 (27 – 51)	36 (25-51)	35 (24-50)	33 (23-49)
Piglets				
All herds	32 (22 – 45)	32 (22 – 45)	32 (22 – 44)	32 (22 – 44)
Small herds (<400)	34 (22 – 50)	34 (22 – 49)	33 (21 – 48)	31 (21 – 47)
Large herds (≥400)	32 (22 – 44)	32 (22 – 44)	32 (22 – 44)	32 (22 – 44)
Sows				
All herds	42 (32 – 55)	41 (31-54)	40 (29-54)	39 (27-53)
Small herds (<200)	42 (33 – 57)	42 (31-55)	41 (29-55)	39 (27-52)
Large herds (≥200)	41 (32 – 53)	40 (30-53)	40 (29-53)	39 (27-54)

Sensitivity of EWS based on mortality rates without a correction for herd size

The EWS as described above was based on mortality incidence rates *with* a correction for herd size. The EWS for CSF was also based on mortality rates *without* a correction for herd size (expressed as no. dead pigs/day), because since the 1st of January 2006 the number of pigs present in a herd are unknown. To identify herds with an extremely high pig mortality every report day, a number of relative and absolute thresholds in mortality rates were determined for each herd type in the same way as determined for the EWS based on mortality incidence rates. No distinction was made between small and large herds, because herd size was assumed to be unknown. The comparison of herds with an extremely high pig mortality between the EWS based on mortality incidence rates (=gold standard) and the EWS based on mortality rates showed that the sensitivity of the EWS based on mortality rates was very poor: The sensitivity was only 33.2% (95% CI: 32.5%-33.8%) for finishing herds, 35.4% (95% CI: 34.6%-36.3%) for piglet herds and 39.3% (95% CI: 38.1%-40.6%) for sow herds compared to the EWS based on mortality incidence rates. In addition, when the EWS based on mortality rates was included in the CSF model, the time to detect a CSF outbreak was not reduced as much as when an EWS based on mortality incidence rates was included (results not shown; Backer et al., 2007b).

DISCUSSION

In this study, an EWS for the control of CSF based on mortality data was developed and its value when added to the current surveillance programmes was quantified. Therefore, all pig mortality data from the Dutch rendering plant and data from the monthly clinical inspections (including herd size) over a period from January 2001-December 2005(CSF-free period) were

used and linked based on unique herd identity number (UHI). However, some herds could not be included in the analyses, because their UHI was missing in the mortality database. In addition, the number of dead piglets that are picked up by the Rendac driver are not reliable, because the number of dead piglets are not correctly reported by the herd owner. Especially during the week-ends, more piglets can die after reporting a number of dead piglets and in such situations the total number of dead piglets that are picked up by Rendac driver is larger than the number reported. In addition, piglets are presented in a barrel and the Rendac driver is not allowed to count the number of dead piglets. For a nationwide EWS for CSF based on mortality data the omission of all UHI's and the discrepancy between the number of dead piglets reported by the herd owner and the number of piglets that are picked up by the Rendac driver have to be resolved.

The EWS for CSF as developed in this study is based on mortality incidence rates. A distinction is made between finishing herds, piglet herds and sow herds, because these herd types show different report patterns and mortality incidence rates. In addition, a distinction between small and large herds is relevant, because small herds have higher mortality incidence rates than large herds. However, the number of piglets present in a herd had to be estimated based on the number of sows, because the number of piglets is not registered during the monthly clinical inspections. Therefore, mortality incidence rates calculated for piglet herds are less reliable than for other herd types. In addition, since the 1st of January 2006 herd size (finishing pigs and sows/boars) is not centrally registered anymore in The Netherlands and thus is unknown. Therefore, an EWS based on mortality rates *without* a correction for herd size was developed. The sensitivity for this alternative, when compared to an EWS for CSF based on mortality incidence rates *with* a correction for herd size, was very low varying from 33.2% (95% CI: 32.5%-33.8%) in finishing herds to 39.3% (95%CI: 38.1%-40.6%) in sow herds. Moreover, herd size plays a major role when calculating mortality incidence rates, because when a pig dies in a small herd the incidence rate will be higher than a pig dying in a large herd. In that case, we can use larger thresholds for small herds than for large herds, so that not only small herds will be identified as herds with extremely high pig mortality. Vice versa, when no correction is done for herd size, herds that are reporting a high number of dead pigs (i.e. large herds) are identified as herds with extremely high pig mortality, whereas small herds will be unidentified even if they have high mortality incidence rates.

Using data about on- and off-farm movements from the Identification and Registration Organisation (I&R) could be an alternative for defining small and large herds for finishing herds and sow herds. However, these movements can not be used to estimate the number of piglets present in a herd, because births are not centrally registered in I&R. Therefore, it is important to register herd size in a central database including the number of piglets, to make an EWS for CSF based on mortality data more efficient and reliable.

To identify herds with extremely high pig mortality, a number of absolute and relative thresholds in mortality incidence rates and differences in mortality incidence rates for each herd type and size were determined. Thresholds based on the mortality incidence rate at one specific moment in time are especially relevant in case of an acute high level in pig mortality. Thresholds based on differences in mortality incidence rates are relevant in case of an increase in pig mortality over time. Relative thresholds are determined every report day and therefore seasonal differences in mortality are accounted for. For example, heat stress during the summer period causes high mortality rates in especially sows. Relative thresholds will be set at a higher level during summer time. In addition, absolute thresholds are relevant in case of very high pig mortality levels, because in those situations more herds will be identified as herds with an

extremely high pig mortality than using only relative thresholds. These absolute thresholds could be very important in case of endemic diseases causing high mortality rates in many herds.

Using both relative and absolute thresholds for each herd type and size, the number of finishing herds with extremely high pig mortality was 1.5-2 times higher than the number of sow herds and piglet herds with extreme high pig mortality. This difference was mainly due to the larger number of finishing herds that are active in The Netherlands compared to piglet herds and sow herds. In addition, more large finishing herds and piglet herds were identified as herds with extremely high pig mortality than small finishing herds and piglet herds, because the proportion of large herds was higher for these herd types. The proportion of small and large sow herds was almost the same and therefore small and large sow herds will be equally identified as herds with extreme high pig mortality, because the relative and absolute thresholds were corrected for herd size.

Backer et al. (2007b) concluded that incorporating the EWS based on mortality incidence data would decrease the median time to detect a CSF outbreak in finishing herds and sow herds with two days (assuming 50% mortality caused by CSFV and using the top 1% of herds with the highest mortality incidence rates) compared to current surveillance programmes. In addition, the 2.5% most extreme detection times were shortened by two days in finishing herds and by one day in sow herds. The decrease in time to detect a CSF outbreak was greater in large herds than small herds, probably because more large herds reported dead pigs to Rendac than small herds.

The EWS based on mortality data was not sensitive enough to detect the additional mortality due to CSFV in piglets. One of the reasons for this lack in detection could be that mortality incidence rates in piglet herds are already high and variable under normal circumstances. In addition, as mentioned above, the number of dead piglets reported by the herd owner and the number of piglets present in a herd are uncertain factors and therefore mortality incidence rates are not reliable for piglet herds.

Stegeman et al. (1999) estimated that during the Dutch CSF outbreak in 1997-1998 the average number of secondary outbreaks caused by one infectious herd was 6.8 before the first outbreak was diagnosed when no specific control measures were implemented. They confirmed that the CSF virus transmits very easily to other herds before implementation of control measures, because during this phase the movement of pigs is not restricted. Transmitting CSFV to other herds by animal movements is of minor interest in finishing herds, because those herds are allowed to move pigs off-farm to the slaughterhouse only. However, sow herds are allowed to move pigs off-farm to many other herds: Herds classified as herd type A (breeding herds) are allowed to move pigs off-farm to an unrestricted number of other herds irrespective of herd type, whereas herds classified as herd type B (multiplier herds) are allowed to move pigs off-farm to an unrestricted number of D-herds (finishing herds). In addition, herds classified as herd type C (rearing herds) are allowed to move pigs off-farm to an unrestricted number of A-herds. Moreover, CSF symptoms are in general non-specific, especially at the early onset of the disease, and infected pigs could easily be transported without knowing that these animals are infected with CSFV. Consequently, CSFV can rapidly spread and infect other herds. The size of a CSF epidemic depends on the control measures taken and the number of herds infected at the end of the High Risk Period (HRP). The decrease in time to detect a CSF outbreak when incorporating the EWS based on mortality incidence data could result in a faster response in taking the first control measures and consequently a lower number of herds that will be infected at the end of the HRP, because in general the number of CSF infected herds increases

exponentially after CSFV introduction in the country. Consequently, lower costs will be associated with a CSF epidemic (Klinkenberg et al., 2005).

In conclusion, implementation of the EWS for CSF based on mortality data seemed advantageous, because the median time and the 2.5% extreme detection times to detect a CSF infection will be shortened in sow herds and finishing herds and therefore could decrease the chance of a major CSF outbreak and the costs related with a CSF epidemic. However, it is difficult to know the real additional value of the EWS based on mortality data in practice, because the UHI of some herds was missing in the mortality data and therefore excluded from analyses. In addition, the number of dead piglets that were picked up by Rendac and the number of piglets present in a herd are not reliable. The effectiveness of the EWS based on mortality data might increase when mortality data are complete and when the alertness of the farmer and veterinarian decreases. In addition, mortality data of the Dutch pig industry are objectively gathered and not influenced by the willingness of a farmer to report high pig mortality. When other diseases are causing mortality, the EWS could also have value for detection of those diseases.

ACKNOWLEDGEMENTS

This study was part of a GIQS(Grenzüberschreitende Integrierte Qualitätssicherung e.V.) project called “Risico’s beheersen/Risiken beherrschen” and funded by the EU INTERREG IIIA program of the Euregios Rhine Maas North and Rhine Waal, the Northrhine Westfalian Ministry of Economy and Dutch Ministry of Agriculture, Nature and Food Quality as well as the Province of Limburg.

REFERENCES

Backer, J.A., Hagenaars, W.H.M., Engel, B., de Jong, M.C.M. and van Roermund, H.J.W. (2007a). Epidemiological model of CSF (in English). In Bergevoet et al. Vaccinatie bij varkenspest. Epidemiologische en sociaaleconomische effecten. ASG report ASG07-I00442, 164 pp

Backer, J.A., Swart, W.A.J.M., Middelesch H. and van Roermund, H.J.W. (2007b). De meerwaarde van mortaliteitsgegevens by Early Warning voor KVP. ASG report ASG07-I00587 and AHS report 4080039, 36 pp

Bouma, A., de Smit, A.J., de Jong, M.C.M., de Kluijver, E. P. and Moorman, R. J. M. (2000). Determination of the onset of the herd-immunity induced by the E2 sub-unit vaccine against classical swine fever virus. Vaccine 18, 1374-1381

Dewulf, J., Laevens, H., Koenen, F., Vanderhallen, H., Mintiens, K., Deluyker, H. and de Kruif, A. (2001). An experimental infection with classical swine fever in E2 subunit marker-vaccine vaccinated and in non-vaccinated pigs. Vaccine 19, 475-482

Dewulf, J., Laevens, H., Koenen, F., Mintiens, K. and de Kruif, A. (2001). An experimental infection with classical swine fever virus in pregnant sows: transmission of the virus, course of the disease, antibody response and effect on gestation. J. Vet. Med. B 48, 583-591

Elbers, A.R.W., Bouma, A. and Stegeman, J.A. (2002). Quantitative assessment of clinical signs for the detection of classical swine fever outbreaks during an epidemic. Vet. Microbiol. 85, 323-332

Klinkenberg, D., Everts-van der Wind, A., Graa, E.A.M. and de Jong, M.C.M. (2003). Quantification of the effect of control strategies on classical swine fever epidemics. Math. Biosci. 186, 145-173

Klinkenberg, D., Nielen, M., Mourits, M.C.M. and de Jong, M.C.M. (2005). The effectiveness of classical swine fever surveillance programmes in The Netherlands. Prev. Vet. Med. 67, 19-37

Laevens, H., Koenen, F., Deluyker, H., Berkvens, D. and de Kruif, A. (1998). An experimental infection with classical swine fever virus in weaner pigs. Vet. Quart. 20, 41-45

Mangen, M.-J.J. and Burrell, A.M. (2003). Who gains, who loses? Welfare effects of classical swine fever epidemics in the Netherlands. European Review of Agricultural Economics 30 (2), 125-154.Moennig, V., 1990. Pestivirus: a review. Vet. Microbiol. 23, 35-54

Moormann, R.J.M., Bouma, A., Kramps, J.A., Terpstra, C. and de Smit, H.J. (2000). Development of classical swine fever subunit marker vaccine and companion diagnostic test. Vet. Microbiol. 73, 209-219

Murray, G., McCutcheon, S. (1999). Model framework and principles of emergency management. Rev. Sci. Tech. Off. Int. Epiz. 18, 15-20

Stegeman, A., Elbers, A.R.W., Bouma, A., de Smit, H. and de Jong, M.C.M. (1999). Transmission of classical swine fever virus within herds during the 1997-1998 epidemic in The Netherlands. Prev. Vet. Med. 42, 201-218

EPIDEMIOLOGICAL METHODS

DIAGNOSTIC ASSESSMENT WITHOUT CUT-OFFS: MODELLING BOVINE DIGITAL DERMATITIS INFECTION

W.D. VINK*, W.O. JOHNSON, G. JONES AND N.P. FRENCH

SUMMARY

Bovine digital dermatitis (BDD) is an epidermitis which is a leading cause of infectious lameness. The only recognised diagnostic procedure is foot inspection, which is labour-intensive to carry out, is subjective and has limited diagnostic sensitivity. Serology is a potentially more sensitive indicator of infection. A reliable and repeatable ELISA was used to test animals from a cross-sectional study on 8 representative dairy farms in the UK. No 'Gold Standard' was assumed to exist. To make fullest use of the information inherent in the ELISA and covariate data, a model using Bayesian statistics was developed which did not impose a cut-off but instead estimated a probability of infection for each individual animal. By modelling BDD in this way, a more detailed and informative analysis of the farm-level distribution of infection could be performed. By extending the model to enable predictive inference, the results of this work can be generally applied.

INTRODUCTION

Bovine digital dermatitis (BDD) was first described as a clinical condition in 1974 (Cheli and Mortellaro, 1974). It appears to have been a 'true' emerging disease as no reference has been made to the clinical condition before this time. It is particularly prevalent in housed Holstein-Friesian dairy cows worldwide, and is considered to be a leading cause of infectious lameness. BDD lesions, which typically develop on the plantar epidermis, tend to be highly painful; hence, BDD has been identified as a major welfare concern. Although economic impacts to the dairy industry are difficult to quantify, these are likely to be substantial.

The rapidity of the spread of BDD attests to its contagiousness. However, the multifactorial nature of the disease has impeded understanding of its aetiology and pathogenesis. Current evidence supports a bacterial aetiology, which furthermore is probably polymicrobial: multiple species of Gram negative bacteria have been associated with the characteristic lesions (Edwards et al., 2003). A consistent finding has been the presence in lesion material of numerous spirochetes, which appear to be associated with necrotic changes. These were demonstrated by microscopy and immunohistochemistry to be *Treponema* spp. (Demirkan et al., 1998; Demirkan et al., 1999a). On the basis of this evidence, lesion-associated treponemes are considered to be the primary microbiological agents involved in the aetiology of BDD.

* Daan Vink, EpiCentre, Institute of Veterinary, Animal and Biomedical Sciences, Massey University, P/Bag 11 222, Palmerston North, New Zealand. Email: W.D.Vink@massey.ac.nz

Clinical inspection of the foot is the only recognised diagnostic procedure. The presentation and development of the disease have been clinically well described. However, no uniformly accepted standard of classifying the typical lesions exists, although there have been attempts at characterising these systematically by stage or severity (e.g. Döpfer and Willemen, 1998). Proliferative lesions are usually prominent and easily visible. However, early-stage lesions tend to be characterised by tissue erosion and are more easily missed. The predilection site of the lesions on the distal plantar skin means they are frequently masked from sight - particularly when environmental hygiene is poor. Lifting of the feet is a time- and labour-intensive procedure. Rodriguez-Lainz et al. (1998) performed screening by visual inspection of the feet of 117 standing cows in the milk parlour. Comparing the results to the assumed 'Gold Standard' test - clinical inspection of the lifted feet in a chute - they determined a sensitivity of 0.72 and a specificity of 0.99. The relatively low sensitivity implies that the prevalence of the condition on dairy farms will be underestimated in the absence of statistical adjustment.

No diagnostics based on molecular detection techniques have yet been developed for screening or diagnostic application: firstly because the aetiology has not yet been completely clarified, secondly because microbiological research of the BDD-associated treponemes has been problematic, and thirdly because understanding of the immunology is limited. The use of serology is attractive because taking blood samples is quicker and more convenient to perform than lifting feet, and less invasive than taking lesion biopsies. Furthermore, as a quantitative test, no subjective interpretation of BDD lesions is required. The continuous outcome of a serological assay could be more informative of an animal's BDD status than the binary outcome of clinical foot inspection. An indirect ELISA has been developed using *Treponema* spp. antigen (Walker et al., 1997; Demirkan et al., 1999b; Murray et al., 2002). Treponemal IgG_2 antibody titres were found to be significantly higher in lesion-positive animals. This indicated that serology could be a viable alternative to clinical inspection for screening and diagnostic purposes.

Using outputs of recent microbiological research, the ELISA test was further developed. The serology of BDD was subsequently investigated by applying this test to a cross-sectional study dataset incorporating data on animals' lesion status and other covariate information. Bayesian approaches have been increasingly applied in recent years for serological modelling, particularly for validation in the absence of a 'Gold Standard' (Joseph et al., 1995; Johnson et al., 2001; Branscum et al., 2004; Branscum et al., 2005). Advantages of such approaches are that they are flexible, implicitly allowing for uncertainty about the sensitivity and specificity of the test, and disease prevalence for the sample populations, to be incorporated. Also, they can easily accommodate unobserved variables such as an individual's latent disease status.

MATERIALS AND METHODS

Description of the dataset and lab analysis

The dataset was derived from a population-based cross-sectional study carried out on eight representative commercial dairy farms in North West England and North Wales. All cattle present on the study farms on the dates of sampling were included in the study (n=2215); dates of birth were obtained from farm records. Serum samples were taken of all animals. The foot hygiene score (FHS) of all animals was assessed: each foot was scored on a 4-point scale (1: very clean to 4: very dirty); by summing these, an individual-level FHS was obtained. The BDD lesion status of a randomly-selected subset of animals from all management groups, on all farms (n=609) was determined by visual inspection using a modified borescope (Vink et al., 2007).

Not all feet were inspected due to time constraints. Although lesions were characterised on the basis of size, clinical presentation and location, a binary outcome was used for model development.

As the focus of this paper is on the modelling and interpretation of serology, the development of the test ELISA will only be summarily described here. Briefly, microbiological research resulted in 23 pure BDD-associated *Treponema* spp. cultures, originating mostly from the study area. Phylogenetic analysis of the 16S rRNA gene sequence showed clustering of isolates into three species-level subdivisions (Evans et al., 2007). ELISAs using individual antigens with a selected panel of sera showed evidence of serogroups consistent with the *Treponema* species clusters. In the absence of sufficiently detailed knowledge, it was assumed that all three species had pathogenic significance. A mixture of five antigens representing the three species was prepared. This 'cocktail' ELISA was investigated using the same panel of sera as for the individual ELISAs. The 'cocktail' assay results were consistently as high as any of the individual ELISAs – hence, it was considered to function well as a 'catch-all' test. Repeated testing with the same serum panel enabled the Pearson correlation coefficient and intraclass clustering coefficient, which are respective measures of the strength of association and the reliability of the test, to be calculated at 0.97 and 0.96 respectively, indicating that the ELISA was highly reliable. Hence, this 'cocktail' ELISA was used to analyse the study sera. Samples were tested in duplicate; duplicates showing poor agreement were retested. The duplicate mean was expressed as a percentage of the plate positive reference sample: percentage positive, henceforth denoted as PP. Exploratory analysis was carried out to investigate the serological distributions.

Records with missing ELISA values (due to loss of samples or test inconsistency: n=17) and missing age data (n=40) were removed; no foot hygiene scores were missing. A total of 2159 records remained in the dataset.

Biological assumptions

By exploratory analysis of the data, animals were identified with high titres but no BDD lesions and vice-versa. Assuming that treponemal serology is associated with incidence of BDD, a fundamental distinction between *infection* and *disease* was consequently made. An individual's 'true' infection status was considered to be unobserved; this was designated as a latent variable (Yang and Becker, 1997). On the other hand, an animal's disease status is defined by presence or absence of lesions as determined by visual foot inspection.

Two screening tests were applied: serology and visual inspection. Neither were assumed to be 'Gold Standard' tests but imperfect indicators of the latent infection status. Furthermore, the responses to the two tests were assumed to be conditionally independent. As the test outcomes are determined by different pathophysiological mechanisms, with responses measured on different outcome scales, this assumption seemed reasonable.

Construction of the basic model

Model specification

Typically, validation of a diagnostic test with a continuous outcome involves specifying an appropriate cut-off point. Assuming a degree of overlap of the distributions of the serological results of the disease-positive and disease-negative sub-populations, the decision where to place this cut-off is a compromise between sensitivity and specificity of the test; a relationship that is commonly investigated using ROC curves (Greiner et al., 2000). By dichotomising the continuous outcome of the serological assay, the absolute difference between each serological result and the specified cut-off is ignored. Thus a PP marginally higher than an arbitrary cut-off would be considered equally positive as a very high PP, whereas in reality test outcomes close to the cut-off have a higher likelihood of being misclassified.

An alternative to the convention of dichotomising test outcomes is to interpret the test as arising from a probability distribution, thereby optimising the information inherent in the continuous test outcome. Such an approach has been described by Choi et al. (2006), who used Bayesian statistics to estimate a predictive probability of infection for a given serological outcome. This approach was taken. The model was implemented in OpenBUGS software[1]. Some explanation follows.

Infection was denoted as I, and no infection as $\bar{I}$. I is a dichotomous variable, hence Bernoulli distributed with probability p_i of an animal i being infected, i.e.

$$I_i \sim \text{Bernoulli}(p_i) \quad (1)$$

From analysis of the serological distributions (Figure 1), a log transformation was considered to provide the closest approximation to a binormal distribution. The log of the PP was denoted as S; therefore the data consisted of $\{S_i;\ i = 1,...,2159\}$. The log serological mean was taken as μ and the precision (which is defined as the inverse of the variance, i.e. $1 / \sigma^2$) as τ, i.e.

$$S_i \sim \text{Normal}(\mu_i, \tau_i) \quad (2)$$

From Figure 1b, it was apparent that the distributions differed for the infected and uninfected sub-populations. This was accounted for by letting $\mu_i = \{(1 - I_i) * \mu_1\} + \{I_i * \mu_2\}$ and $\tau_i = \{(1 - I_i) * \tau_1\} + \{I_i * \tau_2\}$. Hence, for the subpopulation of infected animals, $\mu_i = \mu_2$ and for the subpopulation of uninfected animals, $\mu_i = \mu_1$ (and likewise for τ). Separate prior distributions could then be specified for these parameters.

Presence of disease was defined as presence of at least one lesion and denoted as L, absence thereof as $\bar{L}$. Lesion status, like infection status, is binary, hence Bernoulli distributed with probability q_i:

$$L_i \sim \text{Bernoulli}(q_i) \quad (3)$$

[1]OpenBUGS Version 3.0.1, http://mathstat.helsinki.fi/openbugs/ [Consulted May 2007]

The usual procedure now would be to use explanatory variables to model infection status I_i. Subsequently, the relationship between infection status and disease status would be specified by

$$q_i = I_i * \text{Se} + (1 - I_i) * (1 - \text{Sp}) \quad (4)$$

where Se and Sp represent respectively the sensitivity and specificity of lesion inspection as a diagnostic test of infection. However, some rearrangement was required. As infection status was defined as a latent variable, there was no way of assessing whether the explanatory variables were associated with *I*. Statistical analysis had showed these covariates to be strongly associated with clinical BDD (i.e. *L*). Therefore, these data could be applied to model development of disease, which was directly observed for a subset of animals. The explanatory variables which were selected for modelling included FHS, which was considered a proxy for environmental hygiene (denoted as *x*) and age (denoted as *z*). The explanatory variables were standardised and centred, where $xc_i = (x - \bar{x}) / \text{sd}(x)$ and $zc_i = (z - \bar{z}) / \text{sd}(z)$. Thereafter, q_i was modelled by logistic regression, applying the logit link function:

$$\text{logit}(q_i) = \log(q_i / 1 - q_i) = \beta_1 + \beta_2 * xc_i + \beta_3 * zc_i \quad (5)$$

Thus q_i was regarded as the probability of lesions in cows like the *i*th cow, that is with the same FHS and age. Informative prior distributions were induced for the coefficients β_1, β_2 and β_3 (see below).

Applying Bayes' theorem, p_i was now expressed as a function of q_i:

$$p_i = \text{P}(\text{I}_i) = [\text{P}(L_i) * \text{P}(I_i \mid L_i)] + [\text{P}(\bar{L}_i) * \text{P}(I_i \mid \bar{L}_i)] \quad (6)$$

As *I* was defined as the 'true' status and *L* as the 'apparent' status, $\text{P}(I_i \mid L_i)$ represents the predictive value positive (PVP) and $\text{P}(\bar{I}_I \mid \bar{L}_i)$ the predictive value negative (PVN). Substituting:

$$p_i = \text{P}(\text{I}_i) = [\text{P}(L_i) * \text{PVP}] + [(1 - \text{P}(L_i)) * (1 - \text{PVN})] \quad (7)$$

The overall study population prevalence was denoted as Prev, and the sensitivity and specificity of serology as Se and Sp respectively. The relationship between PVP and PVN and Prev, Se and Sp is as follows:

$$\text{PVP} = \frac{\text{Prev} * \text{Se}}{(\text{Prev} * \text{Se}) + (1 - \text{Prev}) * (1 - \text{Sp})} \quad (8)$$

$$\text{PVN} = \frac{(1 - \text{Prev}) * \text{Sp}}{(1 - \text{Prev}) * \text{Sp} + \text{Prev} * (1 - \text{Sp})} \quad (9)$$

In the current example, PVP and PVN are constant if it is assumed that Se and Sp are innate test properties. However, PVP and PVN will differ between populations with different

prevalences. To enable application of the model in all populations, it was therefore necessary to express PVN and PVP in terms of Se, Sp and Prev. When substituting $P(L_i) = q_i$, this gives:

$$p_i = q_i * \left(\frac{\text{Prev} * \text{Se}}{(\text{Prev} * \text{Se}) + (1 - \text{Prev}) * (1 - \text{Sp})} \right) + (1 - q_i) * \left(1 - \left\{ \frac{(1 - \text{Prev}) * \text{Sp}}{(1 - \text{Prev}) * \text{Sp} + \text{Prev} * (1 - \text{Se})} \right\} \right) \quad (10)$$

Construction of the prior probability distributions

As the choice of the prior distributions contributes to the posterior distributions, the use of informative priors leads to better inference compared to 'vague' priors (Dunson, 2001).

From Figure 1b, the distribution of clinically negative animals can be observed to be right-skewed with a heavy tail, whereas the distribution of clinical positives appears to be approximately normal. As the proportion of clinical negatives with high titres is appreciable, Se was expected to be relatively low. On the other hand, there were fewer lesion-positive animals with low titres, therefore Sp was expected to be higher. As estimates for Se and Sp could not be based on existing knowledge, relatively broad priors were specified: for Se, a mean of 0.6 with a 5th percentile of 0.4, and for Sp, 0.9 with a 5th percentile of 0.8. The prior for Se was more diffuse as there was considered to be more uncertainty about this parameter. Using BetaBuster[1], Se ~ Beta(10.90, 7.60) and Sp ~ Beta(42.57, 5.62) were hence specified. A non-informative prior was placed on the prevalence, Prev ~ Beta(1.90, 1.90), which allowed for a range extending from zero to one, while maintaining a belief that it was more likely to be moderate.

Eliciting information to directly construct informative priors for the regression coefficients is difficult. However, such priors can be induced by using data and expert opinion to specify probabilities for the means of different covariate values; this methodology was developed by Bedrick et al. (1996). First, define the probability of L for an animal on a representative farm, with average FHS ($x = \bar{x}$) and average age ($z = \bar{z}$), to be $\tilde{Q}_1$. In this case, $xc = 0$ and $zc = 0$; therefore, $\tilde{Q}_1 = P(L \mid xc = 0, zc = 0)$. Observe that, by substituting into Eq. (5),

$$\beta_1 = \text{logit}(\tilde{Q}_1) \quad (11)$$

Now let $\tilde{Q}_2$ be the corresponding probability of L for an animal with average age and a specific FHS that was 'above average', where this score was given by x, i.e. $\tilde{Q}_2 = P(L \mid x, zc = 0)$. Again substituting into the regression equation,

$$\text{logit}(\tilde{Q}_2) = \beta_1 + \beta_2 * (x - \bar{x}) / \text{sd}(x) \quad (12)$$

As $\beta_1 = \text{logit}(\tilde{Q}_1)$, it is possible to rearrange and solve for β_2:

$$\beta_2 = [logit(\tilde{Q}_2) - logit(\tilde{Q}_1)] * \text{sd}(x) / (x - \bar{x}) \quad (13)$$

[1] BetaBuster, http://www.epi.ucdavis.edu/diagnostictests/betabuster.html [Consulted August 2006]

Finally, consider $\tilde{Q}_3$, the probability of L for animals with average FHS and an 'above average' age, where this score was given by z. Following the same steps as above, $\tilde{Q}_3 = \mathrm{P}(L \mid xc = 0, z)$. Therefore, $\mathrm{logit}(\tilde{Q}_3) = \beta_1 + \beta_3 * (z - \bar{z}) / \mathrm{sd}(z)$ and

$$\beta_3 = [\mathrm{logit}(\tilde{Q}_3) - \mathrm{logit}(\tilde{Q}_1)] * \mathrm{sd}(z) / (z - \bar{z}) \quad (14)$$

Information sources independent of the data, including scientific literature (Murray et al., 1996; Somers et al., 2005; Holzhauer et al., 2006) and the authors' experience on regional dairy farms, were utilised where possible to specify the 'best estimates' for $\tilde{Q}_1$, $\tilde{Q}_2$ and $\tilde{Q}_3$ (Table 1). Higher FHS was assumed to be a risk factor for BDD; on the other hand, lower BDD prevalence in older cows indicated that increasing age exerted a protective effect, possibly due to development of partial immunity. Beta distributions were used, since these provide a rich family with which to specify prior information, and for flexibility and ease of computation (Enøe et al., 2000). These subsequently induced priors on the regression coefficients.

Extending the model for predictive inference

The basic model specified above results in posterior estimates of parameters, which can be applied to investigate the farm-level distribution of infection of the dataset. In particular, the model estimates of $\mathrm{P}(I_i)$ for animal i can be plotted against its corresponding serological test result.

From Eq. (10), it is apparent that the model-based inferences of the probability of infection depend the probability of lesions (which is informed by the covariates FHS and age) and the prevalence, Prev, in the eight study herds. For inferences to be valid for other study populations, the model is extended to enable inference independent of covariate data.

Predictive probability of infection

The objective of predictive inference is to obtain an estimate of the probability of infection for what has been conceptually described by Choi et al. (2006) as 'future' serological values; these are denoted as Sf. Such estimates can be obtained for given lesion status, FHS and age. First, define $f(S_\mathrm{f} \mid I)$ and $f(S_\mathrm{f} \mid \bar{I})$ as the probability densities corresponding to the serological test result in infected and uninfected cattle, respectively:

$$f(S_\mathrm{f} \mid I) = e^{-1/2 * (S_\mathrm{f} - \mu_2)^2 * \tau_2} \text{ and } f(S_\mathrm{f} \mid \bar{I}) = e^{-1/2 * (S_\mathrm{f} - \mu_1)^2 * \tau_1} \quad (15)$$

Now applying Bayes' theorem for animals with unknown lesion status, lesion positives, and lesion negatives, respectively:

$$\mathrm{P}(I \mid S_\mathrm{f}, xc, zc) = \frac{f(S_\mathrm{f} \mid I) * \mathrm{P}(I \mid xc, zc)}{f(S_\mathrm{f} \mid I) * \mathrm{P}(I \mid xc, zc) + f(S_\mathrm{f} \mid \bar{I}) * \mathrm{P}(\bar{I} \mid xc, zc)} \quad (16)$$

$$\mathrm{P}(I \mid L, S_\mathrm{f}, xc, zc) = \frac{f(S_\mathrm{f} \mid I) * \mathrm{P}(I \mid L, xc, zc)}{f(S_\mathrm{f} \mid I) * \mathrm{P}(I \mid L, xc, zc) + f(S_\mathrm{f} \mid \bar{I}) * \mathrm{P}(\bar{I} \mid L, xc, zc)} \tag{17}$$

$$\mathrm{P}(I \mid \bar{L}, S_\mathrm{f}, xc, zc) = \frac{f(S_\mathrm{f} \mid I) * \mathrm{P}(I \mid \bar{L}, xc, zc)}{f(S_\mathrm{f} \mid I) * \mathrm{P}(I \mid \bar{L}, xc, zc) + f(S_\mathrm{f} \mid \bar{I}) * \mathrm{P}(\bar{I} \mid \bar{L}, xc, zc)} \tag{18}$$

Both sides of the equation are free of the covariates *xc* and *zc*, based on biology. Furthermore, P(*I*) = Prev, P(*I* | *L*) = PVP and P(*I* | *L*) = PVN. The predictive probabilities of infection for animals with unknown lesion status, PPI, lesion positives, PPI_L, and lesion negatives, $\mathrm{PPI}_{\bar{L}}$, are then obtained by substitution:

$$\mathrm{PPI} = \mathrm{P}(I \mid S_\mathrm{f}) = \frac{f(S_\mathrm{f} \mid I) * \mathrm{Prev}}{f(S_\mathrm{f} \mid I) * \mathrm{Prev} + f(S_\mathrm{f} \mid \bar{I}) * (1 - \mathrm{Prev})} \tag{19}$$

$$\mathrm{PPI}_L = \mathrm{P}(I \mid L, S_\mathrm{f}) = \frac{f(S_\mathrm{f} \mid I) * \mathrm{PVP}}{f(S_\mathrm{f} \mid I) * \mathrm{PVP} + f(S_\mathrm{f} \mid \bar{I}) * (1 - \mathrm{PVP})} \tag{20}$$

$$\mathrm{PPI}_{\bar{L}} = \mathrm{P}(I \mid \bar{L}, S_\mathrm{f}) = \frac{f(S_\mathrm{f} \mid I) * (1 - \mathrm{PVN})}{f(S_\mathrm{f} \mid I) * (1 - \mathrm{PVN}) + f(S_\mathrm{f} \mid \bar{I}) * \mathrm{PVN}} \tag{21}$$

Observe the lack of dependence on FHS and age in the formulae. All parameters have been specified as priors except Sf. To make inferences using Eqns. (19) to (21), a grid of untransformed serology values, covering the full range of PP outcomes, is defined. The specified serology values are transformed in the computer code (as the model is based on transformed values) to obtain Sf. By running the model, posterior mean estimates are then obtained for PPI, PPI_L and $\mathrm{PPI}_{\bar{L}}$; these are subsequently plotted against the grid of serology values.

It is also relevant to make inferences about the predictive probabilities of infection when the lesion status is unknown but covariate data are available, as is the case for most of our data. To obtain plots of the predictive probabilities of infection over the grid of serology scores, a set of fixed FHS and age combinations is firstly defined: animals of three ages (1, 3.5 and 7 years) were defined and, for each age, the corresponding estimated 'typical' FHS values for clean, average and dirty feet. Centred values of these covariate combinations (xc_f and zc_f) are derived from the data, resulting in $\{(xc_{\mathrm{f}j}, zc_{\mathrm{f}j}): j = 1,...,9\}$. Next, the corresponding probability of lesions q_f is estimated, using the model in Eq. (5):

$$\mathrm{logit}(q_{\mathrm{f}j}) = \beta_1 + \beta_2 * xc_{\mathrm{f}j} + \beta_3 * zc_{\mathrm{f}j} \tag{22}$$

Estimates are thereby obtained of $\mathrm{P}(L \mid S_\mathrm{f}, xc_\mathrm{f}, zc_\mathrm{f})$. Subsequently, further application of the law of total probability and substituting gives

$$
\begin{aligned}
\mathrm{P}(I \mid S_{\mathrm{f}}, xc_{\mathrm{f}}, zc_{\mathrm{f}}) &= \mathrm{P}(I \mid L, S_{\mathrm{f}}, xc_{\mathrm{f}}, zc_{\mathrm{f}}) * \mathrm{P}(L \mid S_{\mathrm{f}}, xc_{\mathrm{f}}, zc_{\mathrm{f}}) \\
&\quad + \mathrm{P}(I \mid \bar{L}, S_{\mathrm{f}}, xc_{\mathrm{f}}, zc_{\mathrm{f}}) * \mathrm{P}(\bar{L} \mid S_{\mathrm{f}}, xc_{\mathrm{f}}, zc_{\mathrm{f}}) \\
&= \mathrm{PPI}_{L} * q_{\mathrm{f}\,j} + \mathrm{PPI}_{\bar{L}} * (1 - q_{\mathrm{f}\,j}) \qquad (23)
\end{aligned}
$$

The posterior mean predictive probability of infection is then plotted against the serology values, resulting in a set of 9 plots of the different covariate combinations.

Diagnostic application of the ELISA

Up to this point, the predictive probabilities of BDD infection have been estimated. The results show this is appropriate for epidemiological investigation; however, as such an outcome is indicative of a latent state. It is less informative for the purpose of individual assessment of disease status, i.e. for diagnostic evaluation. In this case, it would be desirable to use serology as an alternative to lesion detection by direct inspection. This can be simply achieved within the existing model framework.

Following the notation introduced above, the goal is to obtain the predictive probability of having one or more lesions given the serological result. This can be achieved whether covariate data are available or not, by utilising Eqns. (23) and (19) and applying alternative forms of Bayes' theorem:

$$
\begin{aligned}
\mathrm{P}(L \mid S_{\mathrm{f}}, xc_{\mathrm{f}}, zc_{\mathrm{f}}) &= \mathrm{P}(I \mid S_{\mathrm{f}}, xc_{\mathrm{f}}, zc_{\mathrm{f}}) * \mathrm{P}(L \mid I) + \mathrm{P}(\bar{I} \mid S_{\mathrm{f}}, xc_{\mathrm{f}}, zc_{\mathrm{f}}) * \mathrm{P}(L \mid \bar{I}) \\
&= \mathrm{P}(I \mid S_{\mathrm{f}}, xc_{\mathrm{f}}, zc_{\mathrm{f}}) * \mathrm{Se} + \mathrm{P}(\bar{I} \mid S_{\mathrm{f}}, xc_{\mathrm{f}}, zc_{\mathrm{f}}) * (1 - \mathrm{Sp}) \qquad (24)
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{P}(L \mid S_{\mathrm{f}}) &= \mathrm{P}(I \mid S_{\mathrm{f}}) * \mathrm{P}(L \mid I) + \mathrm{P}(\bar{I} \mid S_{\mathrm{f}}) * \mathrm{P}(L \mid \bar{I}) \\
&= \mathrm{P}(I \mid S_{\mathrm{f}}) * \mathrm{Se} + \mathrm{P}(\bar{I} \mid S_{\mathrm{f}}) * (1 - \mathrm{Sp}) \qquad (25)
\end{aligned}
$$

Note that in both cases, the resulting predictive probabilities are bounded by the values of Se and (1 – Sp).

RESULTS

Serological frequency distributions

Figure 1a shows the serological frequency distribution of the whole study population, which suggests a degree of bimodality. A log transformation approximated a bi-normal distribution, which would be represented as a two-dimensional normal distribution in which the negative and positive sub-populations are normally distributed and independent of each other. Figure 1b, showing distributions of the clinically inspected subset, confirms the existence of serological sub-populations. While there is substantial overlap between these, the boxplots in Figure 1c show that the median antibody titre of lesion positive animals is significantly higher than that of lesion negative animals.

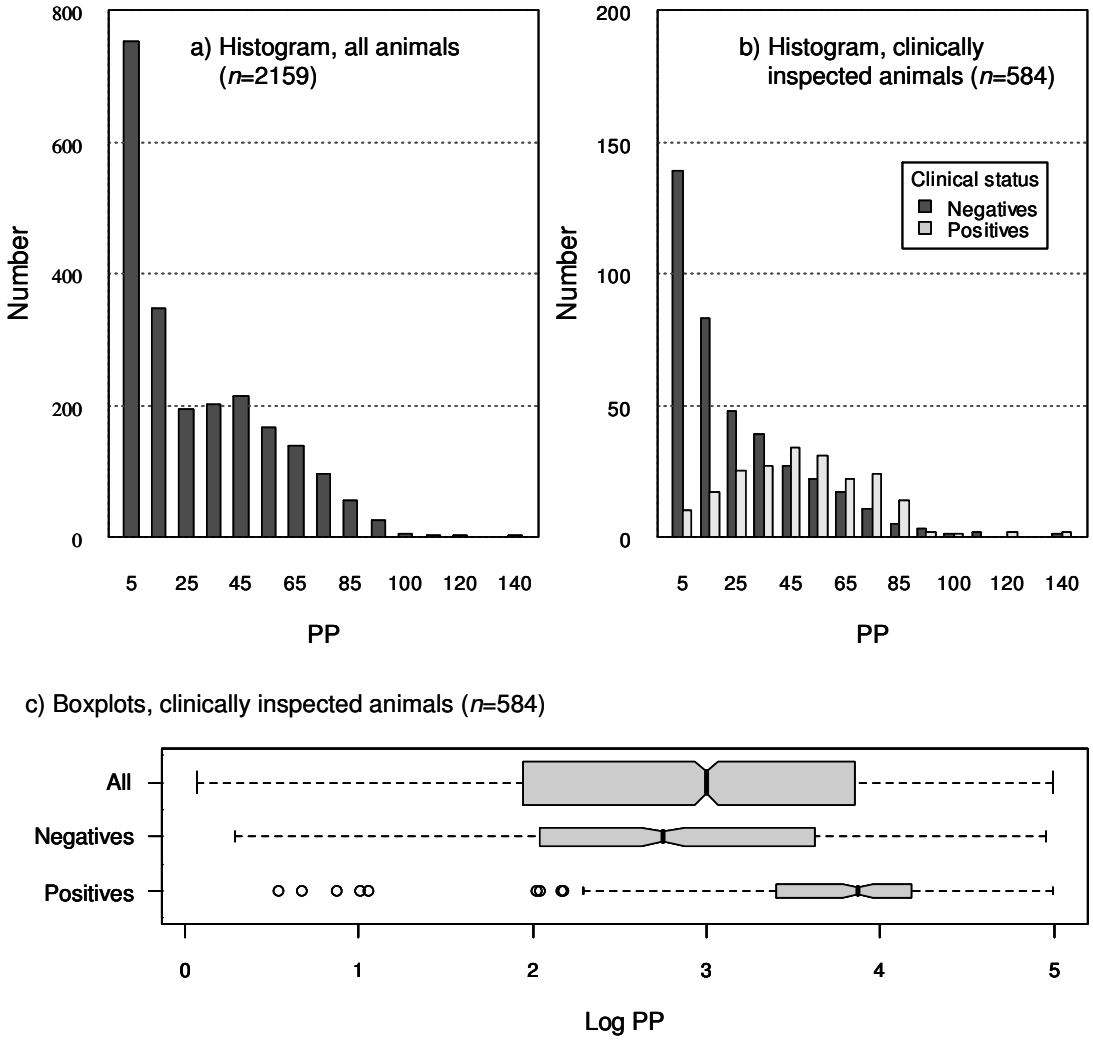


Figure 1. Histograms and notched boxplots for the ELISA of the entire study population and animals inspected for BDD, classified as clinical negatives (n=376) and clinical positives (n=208). The width of the boxplot is proportional to the number of observations; the notches extend to $\pm 1.58 * \left(\mathrm{IQR}/\sqrt{n}\right)$, and represent the approximate 95% confidence interval for the difference in two medians.

Model outputs

Convergence and reliability

The model converged well and consistently gave sensible output. The final model outputs were taken from a run of 20,000 iterations, of which the first 500 were discarded as the burn-in phase. Sensitivity analysis was performed by investigating the effect of systematically varying the priors; the model appeared robust and converged in all cases. The Monte Carlo errors were less than 5% of the standard deviation; autocorrelation was not substantial. Split-sample

analysis, for which the model was run separately on various combinations of data subsets, resulted in highly comparable output.

Summary statistics of the prior and posterior estimates of the key model parameters are given in Table 1. The posterior distributions of the sensitivity, specificity, prevalence and $\tilde{Q}_1$, $\tilde{Q}_2$ and $\tilde{Q}_3$ were narrower than the prior distributions. The posterior mean estimate of prevalence of infection (0.42) was higher than the prevalence of clinical lesions (0.35); this was a consequence of lesion negative animals with high titres. The posterior mean estimate of sensitivity, at 0.76, was better than the prior; likewise, the estimate of the posterior mean specificity (0.95) was higher than the prior.

Table 1: Estimates of model parameters, corresponding prior distributions and posterior summary statistics (20,000 iterations, burn-in of 500 iterations discarded).

	PARAMETER ESTIMATE	5^{TH} / 95^{TH} %ILE	BETA PRIOR DISTRIBUTION	POSTERIOR MEAN	SD	95% PI	MC ERROR
Sensitivity	0.60	0.40	(10.90, 7.60)	0.76	0.08	0.59 – 0.88	0.0031
Specificity	0.90	0.80	(42.57, 5.62)	0.95	0.02	0.90 – 0.98	0.0009
Prevalence	0.50	0.88	(1.90, 1.90)	0.42	0.08	0.28 – 0.58	0.0036
PVN*	-	-	-	0.85	0.02	0.81 – 0.85	0.0010
PVP*	-	-	-	0.92	0.01	0.89 – 0.94	0.0005
$\tilde{Q}_1$	0.25	0.45	(5.41, 14.22)	0.30	0.01	0.27 – 0.33	0.0005
$\tilde{Q}_2$	0.45	0.65	(7.94, 9.49)	0.49	0.02	0.45 – 0.53	0.0008
$\tilde{Q}_3$	0.20	0.40	(4.46, 14.84)	0.44	0.02	0.40 – 0.49	0.0008

* Priors induced by specification of sensitivity, specificity and prevalence, SD: standard deviation; 95% PI: 95% probability interval

Plotting the probability of infection

The model estimate of the probability of infection of each animal, P(I_i), was plotted against its corresponding ELISA result (PP). Lines of best fit were drawn using a smoothing spline. The results, for the entire dataset as well as the stratified subsets, are presented in Figure 2. 81.9% of individuals had a PP <20 or >40. Availability of data on lesion status improved model performance (less noise about the fitted line of lesion positives and negatives, compared to uninspected animals). Compared to the fitted line of the entire dataset, the line of lesion positives is steeper, shifted to the left, and reaches a probability of infection of about 1.0 by a PP of 40; whereas that of lesion negatives is shifted to the right, is less steep, and does not exceed a probability of infection of about 0.85. As lesion negatives with PP >80 were sparse, an edge effect is visible. The line of uninspected animals is highly comparable to that of the entire dataset.

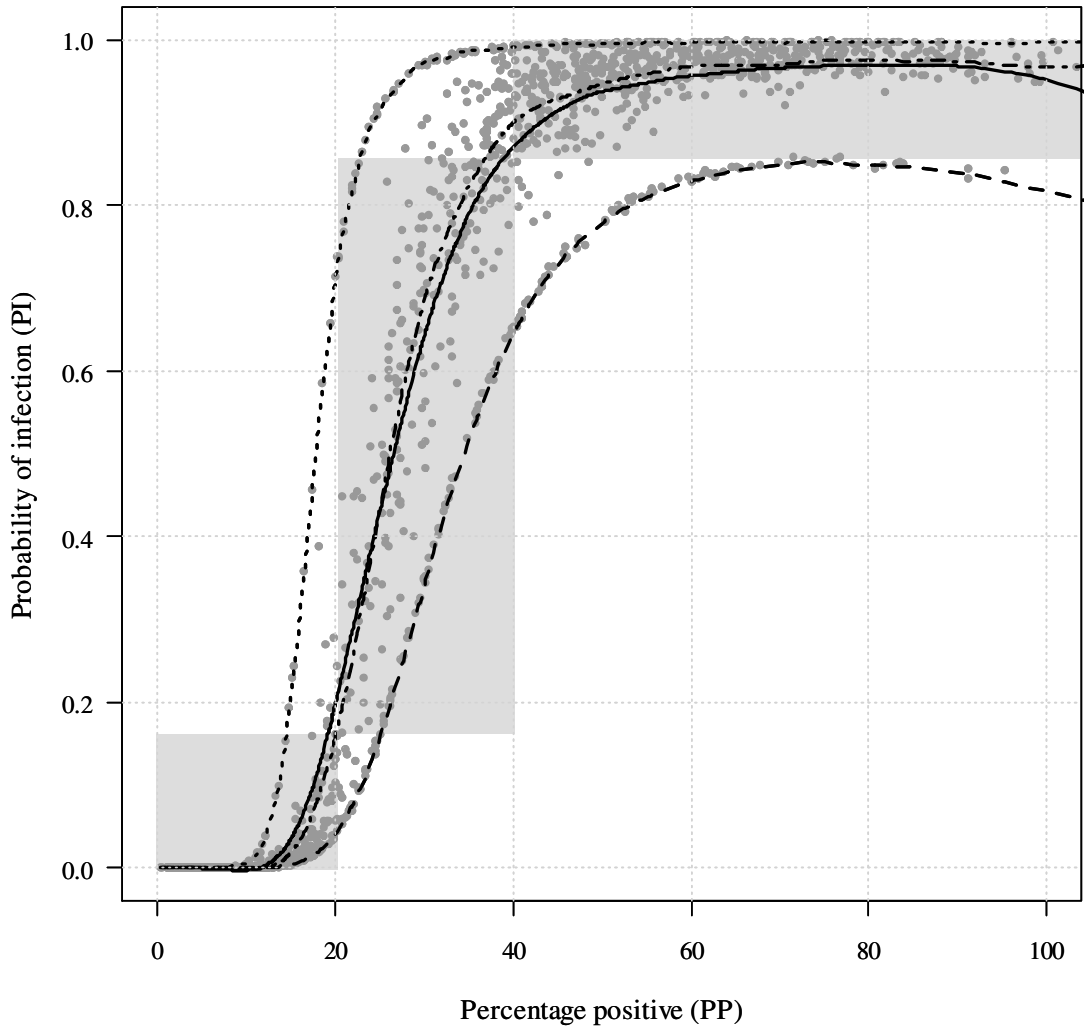


Figure 2. Probability of infection plot as a function of serology, given the data, as estimated by the model. The grey points represent the *P*(*I*) of each animal. Lines were fitted using a smoothing spline: -------- all animals (n=2159); ····· lesion positive animals (n=208); - - - - - lesion negative animals (n=376); – · – · – uninspected animals (*n*=1575). The shaded blocks arbitrarily represent intervals of low *P*(*I*) (PP 0-20), indeterminate *P*(*I*) (PP >20-40) and high *P*(*I*) (PP >40).

Farm-level distribution of infection

The multiple histogram in Figure 3a shows that the model tends to differentiate between uninfected (i.e. low P(*I*)) and infected (i.e. high P(*I*)) animals. The model assigned a P(*I*)<0.10 to 99% of calves and 89% of heifers, resulting in low mean probability of infection in these groups (Figure 3b). Cows tended to be classified as either having low P(*I*) (30% with values of <0.10) or high P(*I*) (42% with values of >0.90). There were few values in the range 0.15-0.85. Mean P(*I*)s with 95% Bayesian credible intervals were 0.09 (0.06-0.12) for calves, 0.28 (0.24-0.33) for heifers and 0.58 (0.57-0.61) for cows.

Of the inspected animals, the relationship between lesion presentation (where this was divided into negative, acute, chronic or regressing lesions) and the probability of infection, as estimated by the model, was investigated. The results are shown in the multiple histogram and boxplots in Figures 3c and 3d.

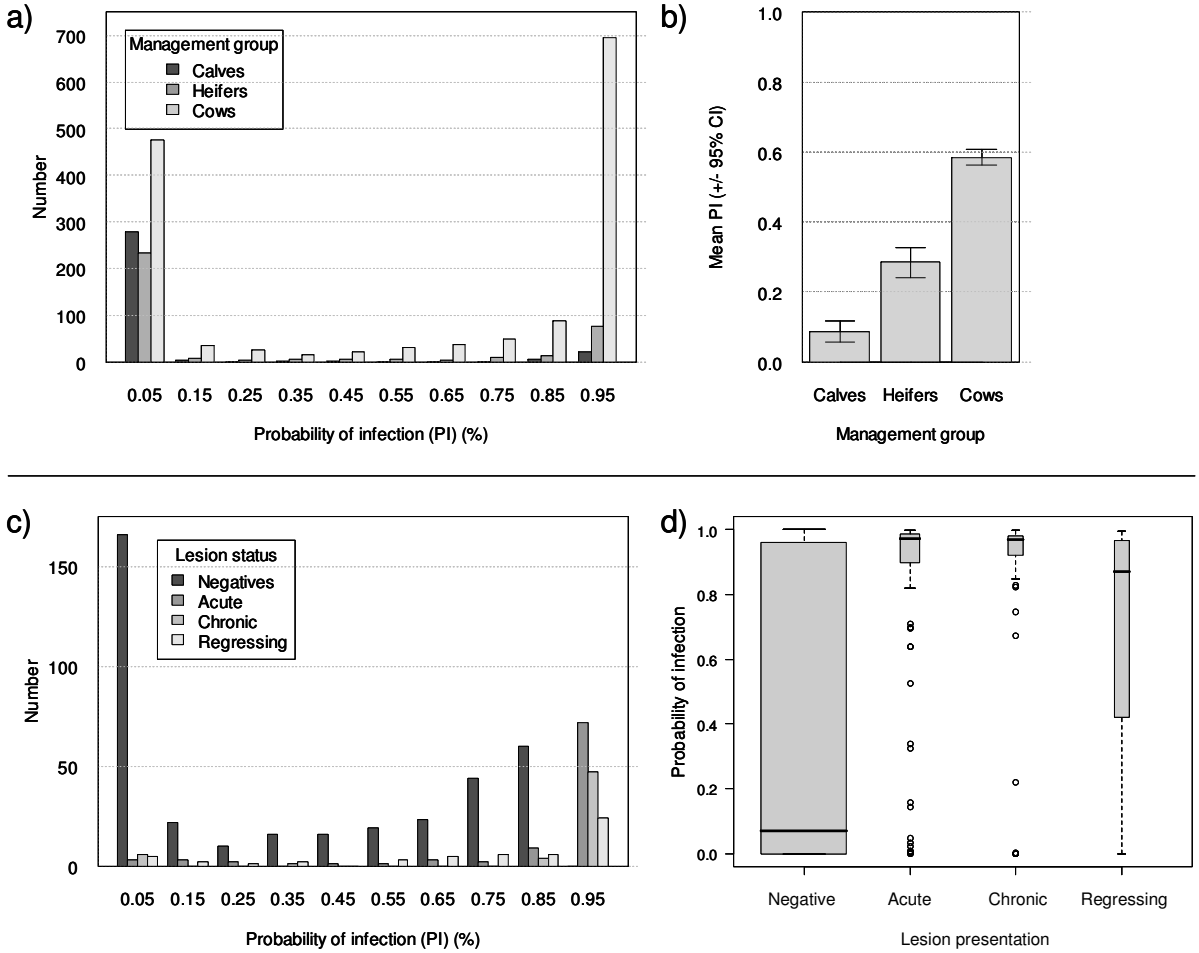


Figure 3. Farm-level distribution of BDD infection as estimated by the model, stratified by management group (top, all animals; n=2159) and lesion status (bottom, inspected animals; n=584). Multiple histograms a) and c) show the probability of infection of animals, ranging from 0 to 99.9% and subdivided into 10 categories. Barplots b) show the mean PI with 95% Bayesian credible intervals. The width of the boxplots d) is proportional to the number of outcomes.

Predictive inference

Figure 4a shows the predictive probability of infection for animals with different lesion status. These are very similar to the P(*I*) plot in Figure 2. Plots of predictive probability of infection for animals of the three ages (1.5, 3 and 7 years) and foot hygiene scores (clean, medium and dirty feet) are shown in Figure 4b. Finally, the predictive probability of having one or more BDD lesions for a given serological test result (in the absence of covariate information) is given in Figure 4c.

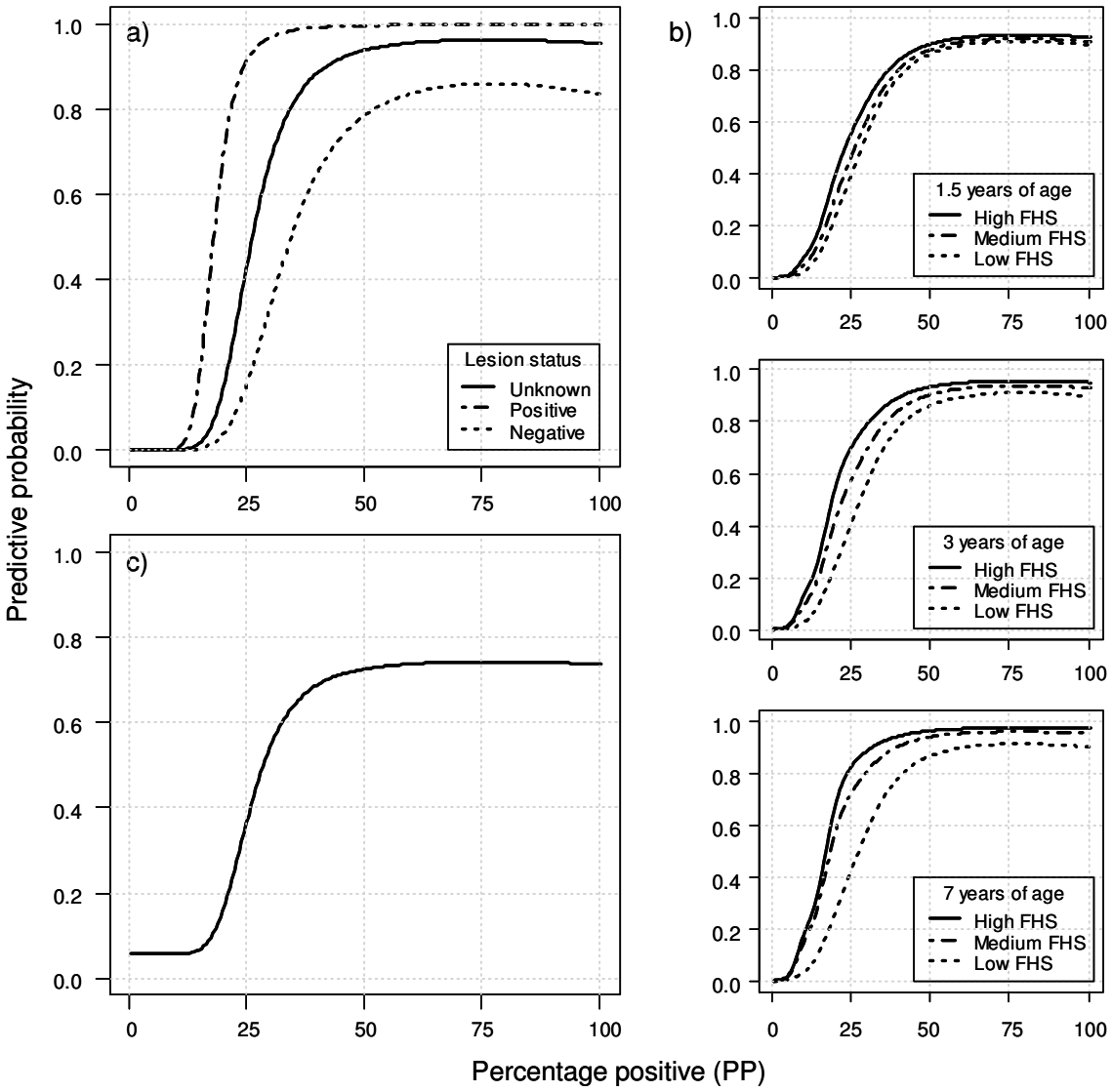


Figure 4. On the basis of a grid of 'future' serological results, S_f: a) Predictive probability of infection for different BDD lesion status, b) Predictive probability of infection for 9 different combinations of age and FHS categories, c) Predictive probability of lesions, in the absence of covariate information.

DISCUSSION

Bovine digital dermatitis remains a relatively poorly understood disease. While the microbial aetiology and the pathogenesis have not yet been definitively elucidated, it is clear that BDD is a complex and multifactorial condition. The 'cocktail' ELISA was developed primarily as an epidemiological tool. By utilising existing knowledge, supplemented by a few biologically plausible assumptions, and applying recent and innovative analytical techniques, the farm-level distribution of infection could be described in more detail than would have been the case if foot inspection alone were performed. Furthermore, assuming that the model was correctly specified and the priors were defined using representative information, the outcomes can be used for interpretation of ELISA results from other dairy cows, i.e. for predictive inference.

A key assumption underpinning the validity of serology is that lesion-associated *Treponema* spp. are implicitly involved in the causation of BDD, and that the IgG_2 antibodies measured by the ELISA are raised as a humoral response to treponemes in lesions. Analysis of the cross-sectional study dataset showed that animals with BDD lesions had significantly higher antibody titres than lesion-negative animals. While the serological distributions showed a degree of bi-modality, there was clearly substantial overlap between the clinical negative and clinical positive sub-populations - in particular, there were relatively many lesion-negative animals with high titres. If the serological response is effected by exposure to BDD-associated treponemes, it follows that exposure does not necessarily result in development of lesions. It therefore seemed reasonable to assume that for modelling BDD, infection status could not be defined by presence or absence of lesions. Similarly, the complexity of the serological response showed that lesion inspection should not be considered as the 'Gold Standard' diagnostic test for BDD infection. For the purposes of epidemiological modelling and investigation, infection status was consequently defined as a latent variable, and a 'Gold Standard' test for its detection was considered not to exist. The substantial overlap between serological distributions of the sub-populations implies that dichotomising the diagnostic test would lead to low sensitivity and/or specificity, and would also lead to loss of information. Hence the Bayesian methodology developed by Choi et al. (2006) was applied, which enables predictive inferences to be made about the probability of infection on the basis of the serological results alone. As an estimation of the probability of infection is less useful in the diagnostic setting (in which case it is desirable to use serology for making an inference about the disease status, i.e. as a proxy for foot inspection), the model was extended to enable estimation of the probability of having one or more BDD lesions.

The P(*I*) plot showed a sigmoid curve with a steeper gradient than might be expected given the substantial overlap of serological distributions between lesion negative and lesion positive sub-populations. This is a desirable characteristic, as the 'indeterminate' interval of serological outcomes (i.e. roughly from PP 20 to 40) was limited. In this range, a small increment of the PP resulted in a large difference in the probability of infection. It is likely that the incorporation of covariate information on foot hygiene score and age, as well as the construction of informative priors by utilising the literature and the authors' experience on dairy farms in the study region, contributed to this.

The curve for lesion positive animals was shifted to the left and had a higher gradient than for lesion negative animals. In other words, a given serology result yielded a higher P(*I*) for lesion-positive animals: such animals with a PP of 20 were estimated by the model to be roughly 80% likely of being infected, versus a corresponding P(*I*) of about 5% for lesion negatives. Lesion negative animals never attained a probability of infection in excess of 85%, regardless how high the PP. The presence of an edge effect for lesion negatives is not surprising, since there were few of these animals with a PP over 90. Comparing the curves for inspected and uninspected animals, it is clear that lesion data improved the precision of the model estimates (there was very little 'noise' in the P(*I*) plots). The fitted line of the uninspected animals was close to that of the overall line, which would be expected.

The individual P(*I*) estimates in the dataset were then used to investigate farm-level distribution of BDD infection. Stratifying by management group, most young stock were unlikely to be infected, whereas cows tended to be discriminately classified into infected or uninfected states. Compared to the underlying serological distribution of the dataset, the model achieves a much clearer division between infected and uninfected animals. Investigating the subset of inspected animals, several striking features emerged. While roughly half of lesion

negative animals had a P(*I*) of <0.10, a substantial number had P(*I*)s up to 0.85, and there was a suggestion of bi-modality. In the corresponding boxplot, this was reflected by the wide upper quartile. A similar, albeit reversed and less obvious, pattern emerges for animals with regressing lesions: while there is an obvious peak at high P(*I*), half of these animals have a P(*I*) of <0.8; the boxplot shows a wide lower quartile. Animals with acute and chronic lesions almost all resulted in a P(*I*) ≥0.80; these categories had high median P(*I*)s with a small interquartile range and few lower outliers, i.e. few animals with these lesions were considered to be uninfected.

The predictive probability of infection for lesion-positive and lesion-negative animals closely mirrored the P(*I*) plot. Generating such plots for different covariate combinations (i.e. different ages and foot hygiene scores) revealed several noteworthy features. For the young animals (1.5 years of age), there was no association between foot hygiene score and the predictive probability of infection; the resulting plot closely resembled that of lesion negatives. The corresponding plots for the cows (3 and 7 years of age) showed that foot hygiene score is a risk factor for BDD (high FHS resulting in a higher predictive probability of infection). Furthermore, age appeared to be associated with BDD, as the curves for the older cows were shifted to the left compared to younger cows.

Although the development of the 'cocktail' ELISA was principally motivated to advance the level of understanding of the farm-level distribution of BDD, it was also interesting to investigate its practical potential for diagnostic application, i.e. to use serology as an alternative for lesion detection by foot inspection. Estimation of the predictive probability of lesions, given a serological outcome, could easily be included in the model. While serology is unlikely to be a substitute for clinical inspection, a scenario could be envisaged in which this could be useful, for example to inform decisions on herd-level intervention strategies (such as footbathing). The plot of predictive probability of lesions showed that serology is less suitable for diagnostic purposes, as it is constrained by Se as the upper bound, i.e. approximately 0.75. Thus the limited sensitivity of the test, which can be attributed to the substantial number of lesion-negative animals with high serology, restricts its applicability. Indeed, this was the underlying reason for assuming the infection status as a latent variable; the model outputs thereby again confirm that modelling infection in such a way was appropriate.

ACKNOWLEDGEMENTS

We would like to acknowledge the participation and assistance of the participating farmers in the field studies. This study was funded by Defra.

REFERENCES

Bedrick, E.J., Christensen, R., and Johnson, W.O. (1996). A new perspective on priors for generalized linear models. J. Am. Stat. Ass. 91: 1450–1460.

Branscum, A.J., Gardner, I.A., and Johnson, W.O. (2004). Bayesian modeling of animal- and herd-level prevalences. Prev. Vet. Med. 66: 101–112.

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.

Cheli, R., and Mortellaro, C.M. (1974). Digital dermatitis in cattle. Proceedings of the 8th International Meeting on Diseases of Cattle, Milan, Italy, pp. 208–213.

Choi, Y.K., Johnson, W.O., and Thurmond, M.C. (2006). Diagnosis using predictive probabilities without cut-offs. Stat. Med. 25: 699–717.

Demirkan, I., Carter, S.D., Murray, R.D., Blowey, R.W., and Woodward, M.J. (1998). The frequent detection of a treponeme in bovine digital dermatitis by immunocytochemistry and polymerase chain reaction. Vet. Microbiol. 60: 285–292.

Demirkan, I., Carter, S.D., Hart, C.A., and Woodward, M.J. (1999a). Isolation and cultivation of a spirochaete from bovine digital dermatitis. Vet. Rec. 145: 497–8.

Demirkan, I., Walker, R.L., Murray, R.D., Blowey, R.W., and Carter, S.D. (1999b). Serological evidence of spirochaetal infections associated with digital dermatitis in dairy cattle. Vet. J. 157: 69–77.

Döpfer, D., and Willemen, M.A. (1998). Standardisation of infectious claw diseases. Proceedings of the 10th International Symposium on Lameness in Ruminants, Lucerne, Switzerland, pp. 244–254.

Dunson, D.B. (2001). Commentary: practical advantages of Bayesian analysis of epidemiologic data. Am. J. Epidemiol. 153: 1222–1226.

Edwards, A.M., Dymock, D., and Jenkinson, H.F. (2003). From hoof to tooth: treponemes in tissue-destructive diseases. J. Appl. Microbiol. 94: 767–780.

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.

Evans, N., Brown, J., Demirkan, I., Murray, R.D., Vink, W.D., Blowey, R.W., Hart, C.A., and Carter, S.D. (2007). Novel spirochetes isolated from Bovine Digital Dermatitis lesions. Emerg. Infect. Dis. In press.

Greiner, M., Pfeiffer, D., and Smith, R.D. 2000. Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. Prev. Vet. Med. 45: 23–42.

Holzhauer, M., Hardenberg, C., Bartels, C.J., and Frankena, K. (2006). Herd- and cow-level prevalence of digital dermatitis in the Netherlands and associated risk factors. J. Dairy Sci. 89: 580–588.

Johnson, W.O., Gastwirth, J.L., and Pearson, L.M. (2001). Screening without a 'Gold Standard': the Hui-Walter paradigm revisited. Am. J. Epidemiol. 153: 921–924.

Joseph, L., Gyorkos, T.W., and Coupal, L. (1995). Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. Am. J. Epidemiol. 141: 263–272.

Laven, R.A., and Lawrence, K.R. (2006). An evaluation of the seasonality of veterinary treatments for lameness in UK dairy cattle. J. Dairy Sci. 89: 3858 – 3865.

Murray, R.D., Downham, D.Y., Clarkson, M.J., Faull, W.B., Hughes, J.W., Manson, F.J., Merritt, J.B., Russell, W.B., Sutherst, J.E., and Ward, W.R. (1996). Epidemiology of lameness in dairy cattle: description and analysis of foot lesions. Vet. Rec. 138: 586–591.

Murray, R.D., Downham, D.Y., Demirkan, I., and Carter, S.D. (2002). Some relationships between spirochaete infections and digital dermatitis in four UK dairy herds. Res. Vet. Sci. 73: 223–30.

Rodriguez-Lainz, A., Melendez-Retamal, P., Hird, D.W., and Read, D.H. (1998). Papillomatous digital dermatitis in Chilean dairies and evaluation of a screening method. Prev. Vet. Med. 37: 197–207.

Somers, J. G. C.J., Frankena, K., Noordhuizen-Stassen, E.N., and Metz, J.H.M. (2005). Risk factors for digital dermatitis in dairy cows kept in cubicle houses in The Netherlands. Prev. Vet. Med. 71: 11–21.

Vink, W.D., Murray, R.D., Carter, S.D., and French, N.P. (2007). Diagnostic assessment of a borescope for bovine digital dermatitis. Vet. J. Submitted.

Walker, R.L., Read, D.H., Loretz, K.J., Hird, D.W., and Berry, S.L. (1997). Humoral response of dairy cattle to spirochetes isolated from papillomatous digital dermatitis lesions. Am. J. Vet. Res. 58: 744–748.

Yang, I. and Becker, M.P. (1997). Latent variable modeling of diagnostic accuracy. Biometrics 53: 948–958.

APPENDIX: MODEL CODE

```
Model
{
for(i in 1:N){
       S[i] ~ dnorm(mmu[i],ttau[i])
              mmu[i] <- ((1 - I[i]) * mu[1]) + (I[i] * mu[2])
              ttau[i] <- ((1 - I[i]) * tau[1]) + (I[i] * tau[2])
       I[i] ~ dbern(p[i])
       p[i]<- (L[i] * PVP) + ((1 - L[i]) * (1 - PVN))
       L[i] ~ dbern(q[i])
       logit(q[i]) <- b[1] + (b[2] * xc[i]) + (b[3] * zc[i])
                     }

for (j in 1:2){
       tau[j] ~ dgamma(0.1,0.1)
       mu[j] ~ dnorm(c[j],d[j])
                     }

# PREDICTIVE PREVALENCES

# "outside loop" computes estimated predictive probabilities of infection and #
lesions over grid of serology values, where m represents the number of rows

for (k in 1:m){
       PPI[k] <- (fSI[k] * Prev) / ((fSI[k] * Prev)
                     + (fSbarI[k] * (1 - Prev)))
       PPI_L[k] <- (fSI[k] * PVP) / ((fSI[k] * PVP)
                     + (fSbarI[k] * (1 - PVP)))
       PPI_barL[k] <- (fSI[k] * (1 - PVN))
                     / ((fSI[k] * (1 - PVN)) + (fSbarI[k] * PVN))
       fSI[k] <- sqrt(tau[2]) * exp(-0.5 * pow(Sr[k] - mu[2], 2) * tau[2])
       fSbarI[k] <- sqrt(tau[1]) * exp(-0.5 * pow(Sr[k] - mu[1], 2) * tau[1])

       PPL[k] <- (PPI[k] * Se) + ((1 - PPI[k]) * (1 - Sp))

# nested inside loop estimates predictive probabilities of infection and
# lesions over grid of covariate values, where r represents the number of
# rows

for (n in 1:r) {
       PPIcov[n,k] <- (PPI_L[k] * qq[n]) + (PPI_barL[k] * (1-qq[n]))
       PPLcov[n,k] <- (PPIcov[n,k] * Se) + ((1 - PPIcov[n,k]) * (1 - Sp))
                     }
             }

for (n in 1:r) {
       logit(qq[n]) <- b[1] + (b[2] * xx[n]) + (b[3] * zz[n])
                     }

# SET PRIORS

PVP <- (Prev * Se) / ((Prev * Se) + (1 - Prev) * (1 - Sp))
PVN <- ((1 - Prev) * Sp) / ((1 - Prev) * Sp + Prev * (1 - Se))
Se ~ dbeta(10.90,7.60)
Sp ~ dbeta(42.57,5.62)
Prev ~ dbeta (1.9,1.9)
b[1] <- logit(qtilde[1])
b[2] <- (logit(qtilde[2]) - b[1]) / ((7 - 5.47) / 1.55)
b[3] <- (logit(qtilde[3]) - b[1]) / ((6 - 3.72) / 2.56)
qtilde[1] ~ dbeta(5.41,14.22)
qtilde[2] ~ dbeta(7.94,9.49)
qtilde[3] ~ dbeta(4.46,14.84)
       }
```

EXPLORATORY MULTIVARIATE ANALYSIS FOR DIFFERENTIATING HUSBANDRY PRACTICES RELEVANT TO DISEASE RISK FOR SMALLHOLDER PIG FARMS IN MADAGASCAR

S. COSTARD*, V. PORPHYRE, S. MESSAD, S. RAKOTONDRAHANTA, H. VIDON, F. ROGER, AND D.U. PFEIFFER

SUMMARY

Multiple factorial analysis (MFA) and hierarchical cluster analysis (HCA) were used to differentiate husbandry practices relevant to contagious disease risk in Malagasy smallholder pig farms. Data from 709 pig farms collected in three study areas were included in the analysis, with variables describing husbandry practices organized in six groups: structure of the farm, animal-contacts, person- and vehicle-contacts, feeding, sanitary aspects, and supplementary variables. The results of the MFA showed that the husbandry practices differed greatly between regions. The HCA identified groups of farms within two regions and suggested variation in professional standards amongst the pig farmers. These differences can be partially explained by variation in: access to professional expertise and technical support, training in farm management and control of diseases, and presence of farmers' associations. Control measures and communication need to be adapted accordingly to reduce the risk of pig diseases in smallholder Malagasy production.

INTRODUCTION

In Madagascar, pig production is very important for smallholder farming communities, but is adversely affected by regular outbreaks of contagious diseases. Husbandry practices are considered to be key factors for the introduction and transmission of contagious swine diseases. An understanding of management and biosecurity practices used in Malagasy smallholder pig farms is necessary to allow the development of recommendations for farmers, with the final aim of reducing risk of disease in smallholder pig production.

Multivariate exploratory analyses have been used in other studies to investigate biosecurity and swine farm management (Hurnik et al., 1994; Rose and Madec, 2002; Boklund et al., 2004; Ribbens et al., in press). The current study makes use of MFA (Escofier and Pages, 1994), which analyses several groups of variables defined for the same observations and shows the relationships existing between the groups of variables. In this study, variables were grouped according to distinct aspects of husbandry practices, aspects which were assumed to have a similar influence on the risk of disease. MFA will therefore result in the identification of the main aspects of management and biosecurity practices differentiating pig farms.

* Solenne Costard, Epidemiology Division, Department of Veterinary Clinical Sciences, The Royal Veterinary College, Hawkshead Lane, North Mymms, Hatfield, Hertfordshire, AL9 7TA, UK. Email: scostard@rvc.ac.uk

MATERIALS AND METHODS

Data collection

A cross-sectional study was conducted from December 2005 to April 2006 in three regions important for pig production in Madagascar. These three study areas were named after their largest city: Ambatondrazaka, Arivonimamo and Marovoay. Stratified multistage sampling was applied to collect data on 200 pig farms per study area, in order to estimate frequency of different aspects of husbandry practices with a 90% confidence level and a precision of 5%, and assuming a frequency of 50% for dichotomous factors. A questionnaire using closed questions was administered by 3 trained interviewers per study area. The following aspects of husbandry practices were investigated: demographics, housing, commercial exchanges, methods of reproduction, contacts with other animals, contacts with persons and vehicles, feeding, animal health management, waste handling, attitude of farmer towards biosecurity.

Statistical analysis

Data entry and data coding were performed using Epi Info 3.3.2, data manipulation using Microsoft Access 2003, and descriptive statistics using STATA 9.2 (Statistical Software: Release 9.2., Stata Corp., College Station, TX, USA). MFA was conducted on data from all pig farms as well as separately for each study area, using the package ade4 (Chessel et al., 2004; Dray et al., 2007) with the statistical software R 2.5.1 (R Development Core Team, 2007).

MFA (Escofier and Pages, 1994) examines the relationships existing between variables separated into different groups. It can be considered as a factor analysis (principal component analysis for quantitative variables, multiple correspondence analysis for qualitative variables) applied to the whole set of variables within which each group is weighted. All elements of the dataset (individuals, variables, groups of variables) are represented graphically in a Euclidean space. The principle of factor analysis is to define projections – or factors – representing the optimal summary of the relationships between variable categories, i.e. differentiate between them as much as possible. A factor is therefore a linear combination of the variables and is characterised by its eigenvalue, which indicates the variance - or inertia - of the dataset it captures. The first factor is the projection for which the variance is maximal, and each consecutive factor is defined so that it captures the variance not explained by the previous factor. Factor analysis can include both active variables, used to calculate factors, and supplementary variables, not used to define factors but projected on these factors. In a MFA, factor analyses are first performed independently for each group of variables. Normalization is then conducted via the division of individual scores by the square root of the first eigenvalue, in order to make the different groups of variables comparable in a global analysis. A factor analysis is finally performed on the merged dataset (obtained by juxtaposing the individual normalized datasets), where each group of variables has an equal *a priori* influence in the global analysis. By setting up common factors for both variables and groups of variables, MFA takes into account the heterogeneity of groups of variables in terms of biological meaning, and allows the identification of the main variables and groups of variables that differentiate between the individuals.

HCA (Everitt, 1974) was then conducted on pig farms' MFA scores, using Ward's criteria for linkage. The variable categories significantly associated with each group of farms were used to describe their characteristics. This was achieved by calculating test values, which measure the

distance between the within-group value and the overall value for each variable category. (Morineau, 1984)

RESULTS

Data on husbandry practices were collected from a total of 709 pig farms: 272 in Ambatondrazaka, 233 in Arivonimamo and 204 in Marovoay. A total of 42 categorical variables were included in the analysis. The variables describing husbandry practices with a potential influence on the risk of disease were organized as five groups of active variables (Table 1): structure of the farm (fencing, flooring, roofing, number of pens, presence of pig farms less than 100m away), animal-contacts (type of confinement, presence of other animals on the premises, origin and destination of pigs, exchange of boars for natural service), person- and vehicle-contacts (people on the farm undertaking other activities linked to the pig production sector, stakeholders from the pig production sector allowed onto farm, vehicles allowed onto premises), feeding (types and origin of feeds) and sanitary aspects (animal health management, control of insects and rodents, cleaning and disinfection of equipment, management of manure, slaughtering of pigs on the premises). The variables describing the production systems and considered as not influencing the risk of disease were included in a group of supplementary variables (Table 2).

Multiple factor analysis

The cumulative percentage of inertia of the first two factors selected for the four MFAs performed separately on all observations, farms from Ambatondrazaka, Arivonimamo and Marovoay were 19.2%, 22.5%, 36.5% and 19.3%, respectively. The global display of the groups of variables included in the MFAs (Fig. 1) gave an indication of their importance for differentiating between pig farms: the larger their inertia on the factors 1 and 2, the more they differentiate between pig farms. Figure 1a showed that overall, the groups of variables differentiating between pig farms were: structure of the farm, sanitary aspects, feeding and animal contacts. For Ambatondrazaka, Fig. 1b showed that the inertia of the five groups of variables on factors 1 and 2 were low, and therefore pig farms were poorly differentiated. The global display for Arivonimamo (Fig. 1c) indicated that pig farms were differentiated by the five groups of variables on factor 1. In Marovoay, the groups of farms were differentiated by: feeding, sanitary aspects, animal-contacts and, to a lesser extent, person-contacts (Fig. 1d). In addition, Fig.2 showed that groups of farms with similar husbandry practices existed, and this grouping seemed related to the study areas. These results suggested that the main husbandry practices differed between regions.

Table 1. Description and frequencies of the main variables included in the MFA as active variables

VARIABLE	CATEGORY	FREQUENCY OF THE GIVEN CATEGORY (%)			
		ALL FARMS	AMB	ARV	MRV
Type of enclosure	Fence	65.7	92.3	7.7	96.6
	Wall	34.3	7.7	92.3	3.4
Flooring	Mud or sand	44.7	24.3	48.1	68.1
	Wood duckboard	30.3	63.6	1.3	19.1
	Cement	25.0	12.1	50.6	12.8
Roofing	Thatched roof	76.3	82.7	49.8	98.0
	Tiles or metal sheet roof	15.5	17.3	25.3	2.0
	Pen in house basement	8.2	0.0	24.9	0.0
Manure management	Collected in a septic tank	15.8	8.1	37.8	1.0
	Used as a fertilizer for crops	41.5	50.4	56.7	12.3
	Disposed of nearby to premises	27.6	20.2	0.4	68.6
	Disposed of far from premises	15.1	21.3	5.1	18.1
Health care provided to pigs	Yes	91.0	98.2	97.0	74.5
Insecticide treatment	Yes	11.9	16.2	14.2	3.4
Type of feeds	Fish meal, blood or meat meals	40.8	49.6	0.0	75.5
	Industrial and agricultural by-products (rice straw and grass)	82.8	97.1	53.7	97.1
	Domestic waste	59.2	51.5	48.5	81.9
	Compound feeds	17.6	8.8	33.9	10.8
Origin of pigs	Live animal markets, other farmers	21.7	20.2	42.5	0.0
	Other farmers only	49.8	65.5	18.4	64.7
	Neither live animal market nor other farmers	8.3	10.3	7.3	6.9
	No answer	20.2	4.0	31.8	28.4
Destination of pigs	Live animal markets, traders, butchers or other farmers	10.9	3.3	25.7	3.9
	Traders, butchers or other farmers (not live animal market)	75.9	83.5	67.8	75.0
	No answer	13.2	13.2	6.5	21.1
Type of confinement	Total	80.8	87.1	64.4	91.2
	Partial	19.2	12.9	35.6	8.8
Presence of cattle on the premises	Yes	30.9	10.3	55.4	30.4
Vehicles allowed onto the premises	Yes	62.3	52.6	97.4	35.3

[a] Ambatondrazaka, [b] Arivonimamo, [c] Marovoay

Table 2. Description and frequencies of the variables included in the MFA as supplementary variables

VARIABLE	CATEGORY	FREQUENCY OF THE GIVEN CATEGORY (%)			
		ALL FARMS	AMB[a]	ARV[b]	MRV[c]
Type of farm	Breeding	7.0	61.8	49.8	44.1
	Farrow-to-Finish	40.2	29.0	40.8	54.9
	Finishing	52.8	9.2	9.4	1.0
Breed(s) of pigs	Exotic	33.5	45.2	43.8	6.4
	Local	30.8	17.7	50.6	25.5
	Crossbred	35.7	37.1	5.6	68.1
Presence of a boar	Yes	11.4	8.5	10.3	16.7
Number of sows	0	53.9	64.0	49.8	45.1
	1 - 2	39.2	33.1	38.6	48.0
	> 2	6.9	2.9	11.6	6.9
Number of finishing pigs	0	16.8	18.0	15.0	17.1
	1 - 5	66.3	75.7	55.4	66.2
	> 5	16.9	6.3	29.6	16.7
Number of unweaned pigs	0	70.8	78.7	60.5	72.1
	1 – 10	22.6	20.2	23.6	24.5
	> 10	6.6	1.1	15.9	3.4
Number of pigs sold in 2005	0	70.8	78.7	60.5	72.1
	1 – 10	22.6	20.2	23.6	24.5

[a] Ambatondrazaka, [b] Arivonimamo, [c] Marovoay

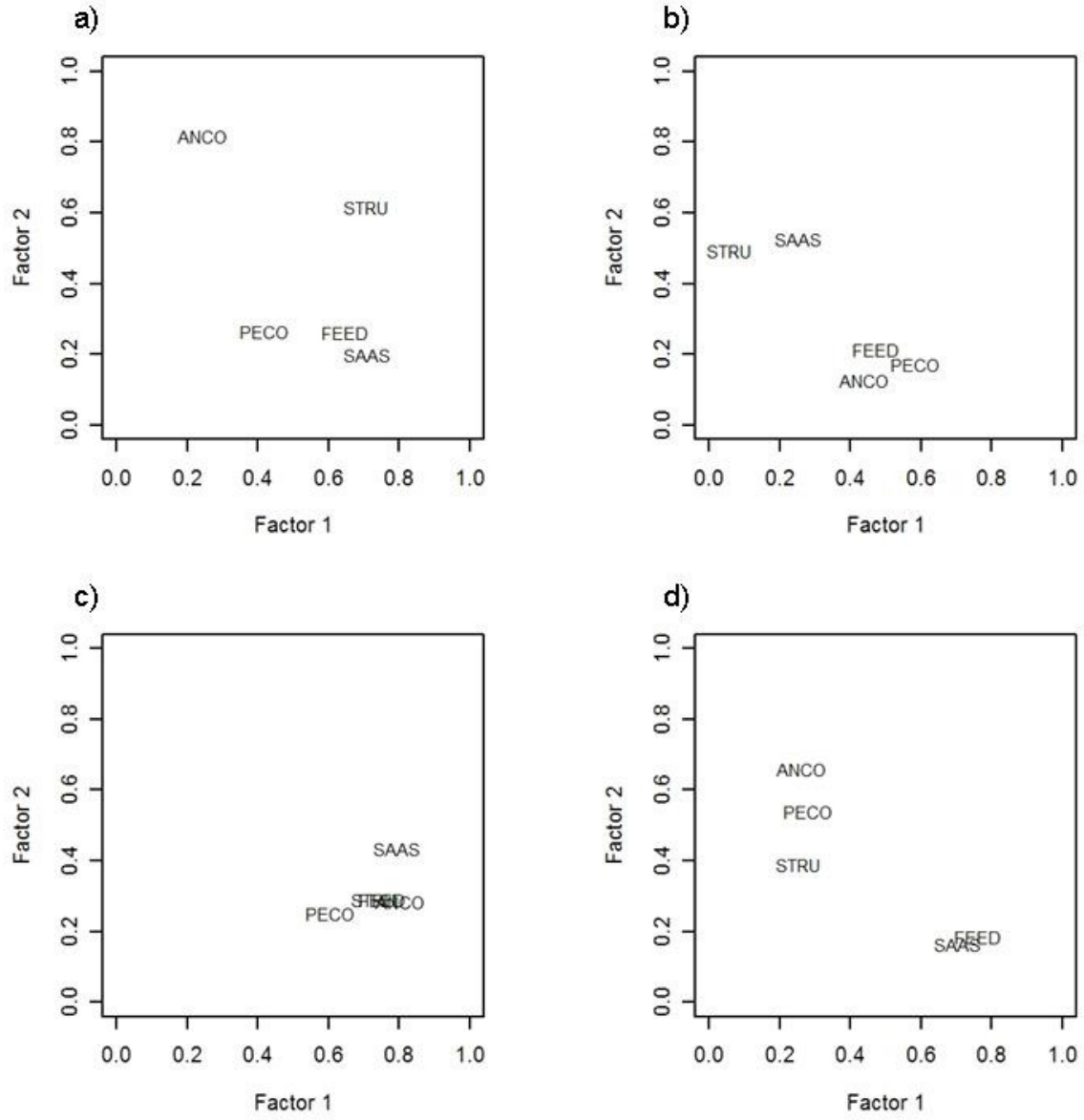


Fig. 1 Global representation of the 5 groups of variables on the first two factors of MFAs for: a) all observations, b) Ambatondrazaka, c) Arivonimamo, d) Marovoay. For each group of variables, their coordinates (between 0 and 1) indicate the inertia explained by the first factor (horizontally) and the second factor (vertically). STRU: structure of the farm; ANCO: animal-contacts; PECO: person- and vehicle-contacts; FEED: feeding; SAAS: sanitary aspects.

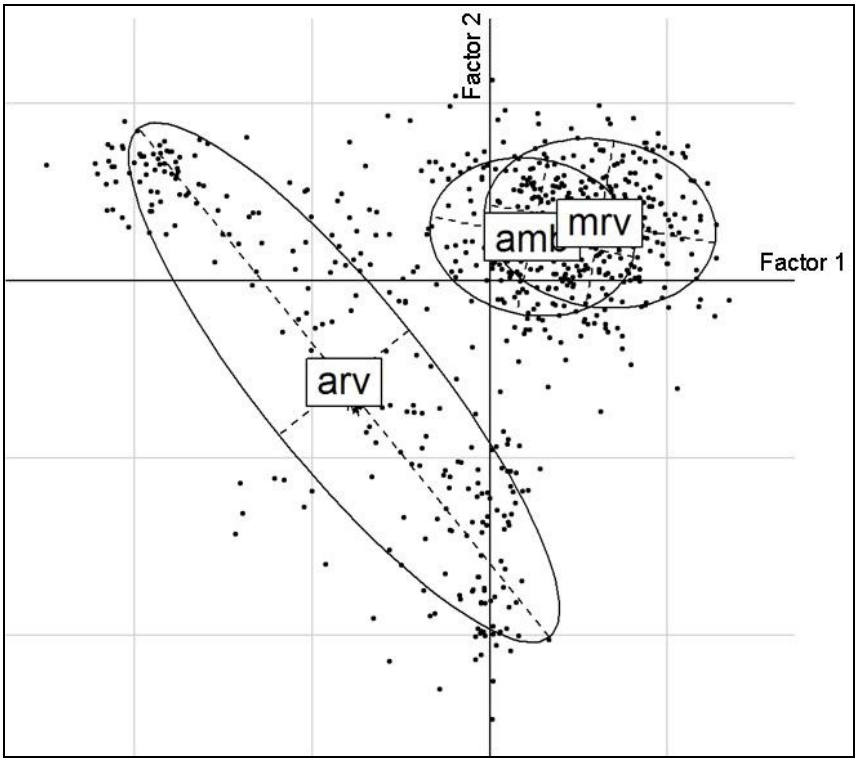


Fig. 2 Representation of all pig farms and study areas (amb: Ambatondrazaka, arv: Arivonimamo, mrv: Marovoay) in relation to the first two factors of the MFA. Points represent pig farms and the distance between them is an indication of their similarity in terms of husbandry practices.

Hierarchical cluster analysis

The HCA carried out on data from each study area resulted in the identification of four clusters of farms in two study areas (Arivonimamo and Marovoay), while no clustering of farms was found in Ambatondrazaka. Although this latter area was not homogeneous in terms of husbandry practices, no clear farm grouping could be differentiated. The main characteristics differentiating the clusters of farms are presented below.

Small farms in Arivonimamo (n=81)

Generally, they were finishing units with pigs of local breed, where less than 10 pigs were sold in 2005. Animals were usually kept in the basement of their owner's house (P=0.021). Compared to other pig farms in Arivonimamo, replacement animals were more often bought on live animal markets (P=0.005); and cattle were more often present on the premises (P=0.026). Pigs were usually fed with domestic waste (P=0.006) and crop by-products (P=0.031). Manure was collected to fertilize crops (P=0.014). People on the farm were more likely to be involved in other activities related to the pig production sector (P=0.014) that in other farms of the area.

Farms with partial confinement in Arivonimamo (n=42)

These were usually finishing farms or breeding units where animals of local breed were kept in pens with mud floor (P=0.029) and mud walls or post-and-rail fences (P<0.001). Compared to other farms in Arivonimamo, more pig owners reported allowing their animals to roam for food (P=0.023), and poultry were less often kept on the premises (P=0.031). Animals were less likely

to receive health care than in the other pig farms of the region (P<0.001), and thus animal health workers were less likely to visit the farms (P=0.034).

Farms with improved biosecurity in Arivonimamo (n=38)

In these farms, pigs of exotic breed or crossbred (P=0.004) were kept permanently in pens made from cement. Compared to other pig farms in Arivonimamo, they were more often situated more than 100m away from other pig farms (P=0.031), fewer vehicles were allowed onto premises (P<0.001), and more owners reported carrying out disinfection (P<0.001). Animals were fed with crop by-products bought from shops (P=0.002) rather than with domestic waste (P=0.040). Commercial exchanges of pigs were made with other farmers or traders (P=0.038).

Large farms in Arivonimamo (n=72)

These farrow-to-finish units with animals from exotic breed had sold more than 10 pigs in 2005. Animals were kept permanently (P=0.025) in buildings with separate pens made from cement (P=0.013). Compared to other pig farms in the region, more owners reported controlling rodents (P=0.032) but not carrying out disinfection or treatments against insects. A larger proportion of farmers fed their animals with compound feeds (P<0.001) bought at markets (P=0.018). Usually replacement animals were not bought from live animal markets (P=0.015). Manure was collected in septic tanks (P<0.001) rather than used as fertilizer for crops. People working on the farm were less likely to undertake other activities linked to the pig production sector (P=0.014) than in other farms of the area.

Small farms in Marovoay (n=45)

For this group of farms, all the test values produced P values higher than 0.05. It was therefore interpreted as the baseline group, describing the average pig farm in Marovoay area. In these small finishing or farrow-to-finish units, pigs of local breeds or crossbed were generally kept in pens with mud floor and post-and-rail fences. These farms were usually situated less than 100m away from other pig farms, and contact between animals from different farms occurred for reproduction or commercial purpose. Poultry and dogs were kept on premises in addition to pigs. Animals were fed with domestic waste, crop by-products and meat or blood meals. Almost no farmers reported carrying out disinfection and treatment against insects and rodents.

Farms with reduced person-contacts in Marovoay (n=39)

Compared to other pig farms in Marovoay, animals were less often sold to traders or other pig farmers (P=0.010). Visits from other stakeholders, and animal health workers in particular (P=0.007), were less frequent than in other farms of the region. Animals were fed with domestic waste and feeds bought on markets (P<0.001) rather than from rice producers.

Farms with numerous person-contacts in Marovoay (n=59)

These farms were farrow-to-finish units with larger herds of crossbred and local breed pigs. Feeds were bought from rice producers and rice plants (P=0.030) as well as from markets. Compared to other farms in Marovoay, a high proportion of boar owners reported lending their animals for natural service (P=0.023). More farmers reported having visits from other pig farmers (P=0.031) and traders (P=0.048), and allowing visitors' vehicles on farms (P=0.021). However, few farmers reported buying animals from other pig farmers, traders or on live animal markets (P=0.046).

Large farms in Marovoay (n=38)

They were farrow-to-finish units with crossbred animals permanently kept in pens of post-and-rail fences or solid walls (P=0.002). Animals were fed with compound feeds (P<0.001) bought in shops (P=0.002) in addition to crop by-products. Compared to other pig farms in the region, a higher proportion of farmers reported carrying out disinfection (P<0.001), control of rodents (P=0.003), and collecting manure in a septic tank (P<0.001).

DISCUSSION

Overall, the husbandry practices reported by farmers indicated a low level of biosecurity. Smallholder pig farming is a traditional farming system where almost no sanitary measures are applied. Many opportunities for contacts with potential sources of infection exist, such as pigs exchanged between farms, for commercial reasons or for natural service, other species of animals and various stakeholders in the pig production sector.

Multiple factor analysis highlighted regional differences in terms of structure of pig farms, sanitary aspects, feeding and animal-contacts. In Ambatondrazaka, pig farms were units with post-and-rail fences, with a high likelihood of contacts with other animals and stakeholders in the pig production sector. They had low levels of cleaning and disinfection of equipment on farm. In Arivonimamo, small traditional farms co-existed with larger farms where biosecurity levels were relatively higher. The former were free-range farms or farms where animals are kept in the basement of the farmer's house, with animals being sold between pig farmers, to traders and to live animal markets. The latter were farms with animals kept permanently in pens made from cement, where manure is collected in a septic tank, buildings are treated against insects, and animals fed with compound feeds rather than crop by-products and domestic waste. In Marovoay, pig farms are traditional farming systems with post-and-rail fences, where animals are less likely to receive health care than in the two other regions.

The hierarchical cluster analysis allowed the identification of groups of farms with distinct husbandry practices in Arivonimamo and Marovoay. The main practices differentiating these groups were: animal-contacts (origin and destination of pigs, type of confinement, presence of other animals on the premises, lending of boar for natural service), feeding (origin of feeds, type of feeds), sanitary aspects (manure management, disinfection, control of rodents, animal health management), person- and vehicle-contacts (visitors allowed onto farm, vehicles allowed onto farm, people on farm undertaking other activities linked to the pig production sector or to wild pigs), structure of the farm (fencing, flooring, roofing, presence of other pig farm(s) less than 100m away).

The heterogeneity observed at the regional level can be explained by differences in ethnicity and culture, climatology, agro-systems and ethnicity (Table 3), but also availability of training and technical advice for farmers between regions. In Ambatondrazaka and Arivonimamo, veterinarians and animal health workers were present in important localities. Training on swine diseases and control measures was provided to farmers, particularly after the introduction of African swine fever in the late 1990s. In Marovoay however, a large number of farmers reported little training and limited access to technical advice. The within-area differences can be explained by variation in: access to professional expertise on swine health, technical support and training of farmers on farm management and diseases, as well as the presence of farmers' associations aimed at improving production. In Arivonimamo for example, small traditional farms with low biosecurity levels are generally situated in remote villages with veterinarians and

animal health workers are less likely to visit these farms and pig owners are less likely to attend training as they need to travel long distances to go to places where the meetings and training courses are held. Considering the professional expertise and technical support available in the main cities of Ambatondrazaka, variation in professional standards amongst pig farmers was expected, as observed in Arivonimamo. The absence of distinct farm groupings in relation to management and biosecurity practices might suggest that farmers have individually adapted their practices in accordance with their own risk perceptions. In both, Arivonimao and Marovoay, farmers' associations are present in some localities. This might be an indication of farmers' willingness to work together in their disease control and pig production and share expertise.

MFA and HCA were well suited for the exploratory analyses of the data collected in this study, partly because they can extract the key information from a large number of correlated variables. MFA allows the analysis of data organised in different groups of variables with distinct biological interpretation. This results in a more intuitive interpretation of the results than with the exploratory methods applied in the studies cited previously, or with the scoring system used by Pinto and Urcelay (2003).

In conclusion, this study provided a description of swine management practices in smallholder pig farms in three regions of Madagascar, and a characterisation of farm types based on husbandry practices relevant to disease risk. In addition, it highlighted out the need to adapt communication with farmers and control measures accordingly in order to reduce the risk of introduction of diseases for pig smallholder communities.

Table 3. Putative explanatory factors for the differences of husbandry practices observed between areas

FACTOR	AMB [a]	ARV [a]	MRV [a]
Climate [b]	Wet Tropical	Temperate	Wet and Dry Tropical
Ethnic group [b]	Sihanaka	Merina	Sakalava
Rice crop management [b, c]	Irrigated Rainfed lowland	Rainfed lowland Upland (slash-and-burn)	Irrigated
Main livestock production [d]	Zebu (agricultural work) Dairy cattle Swine Small ruminants	Zebu (agricultural work) Dairy cattle Swine	Zebu (meat production) Dairy cattle Swine

[a] Ambatondrazaka, [b] Arivonimamo, [c] Marovoay
[b] http://lcweb2.loc.gov/frd/cs/mgtoc.html
[c] http://www.fao.org/AG/Agp/agpc/doc/riceinfo/AFRICA/Madagascar.htm
[d] http://www.ilo.cornell.edu/polbrief/03conv/map3-3.html

ACKNOWLEDGEMENTS

The authors wish to thank the Wellcome Trust for funding this study. We are also very grateful to the farmers and field staff who participated in the study.

REFERENCES

Boklund, A., Alban, L., Mortensen, S., Houe, H. (2004). Biosecurity in 116 Danish fattening swineherds: descriptive results and factor analysis. Prev. Vet. Med. 66, 49-62

Chessel, D., Dufour, A.B., Thioulouse, J. (2004). The ade4 package - I: One-table methods. R News 4, 5-10

Dray, S., Dufour, A.B., Chessel, D. (2007). The ade4 package - II: Two-table and K-table Methods. R News 7, 47-52

Escofier, B., Pages, J. (1994). Multiple Factor Analysis (Afmult Package). Computational Statistics & Data Analysis 18, 121-140

Everitt, B. (1974). Cluster Analysis. Heinemann Educ. Books, London.

Hurnik, D., Dohoo, I.R., Donald, A., Robinson, N.P. (1994). Factor analysis of swine farm management practices on Prince Edward Island. Prev. Vet. Med. 20, 135-146

Morineau, A. (1984). Note sur la caractérisation statistique d'une classe et les valeurs tests. Bulletin Technique du Centre de Statistique et d'Informatique Appliqués 1, 9-12

Pinto, C.J., Urcelay, V.S. (2003). Biosecurity practices on intensive pig production systems in Chile. Prev. Vet. Med. 59, 139-145

R Development Core Team (2007). R: A language and environment for statistical computing (Vienna, Austria, R Foundation for Statistical Computing).

Ribbens, S., Dewulf, J., Koenen, F., Mintiens, K., De Sadeleer, L., de Kruif, A., Maes, D. A survey on biosecurity and management practices in Belgian pig herds. Prev. Vet. Med. (2007), doi:10.1016/j.prevetmed.2007.07.009

Rose, N., Madec, F. (2002). Occurrence of respiratory disease outbreaks in fattening pigs: Relation with the features of a densely and a sparsely populated pig area in France. Veterinary Research 33, 179-190

THE USE OF CRITICAL REVIEWS AND META-ANALYSIS OF DIAGNOSTIC TEST EVALUATIONS – ILLUSTRATED BY DIAGNOSTIC TESTS FOR PARATUBERCULOSIS

N. TOFT* AND S. S. NIELSEN

SUMMARY

This study presents summary receiver operating characteristics (SROC) analysis as a means to synthesise information from diagnostic test evaluation studies. Using data from a review of diagnostic tests for ante-mortem diagnosis of paratuberculosis, SROC and hierarchical SROC (HSROC) analysis were used to estimate overall diagnostic accuracies of antibody ELISAs for bovine paratuberculosis while accounting for covariates such as the target condition used in the test evaluation. The methods gave comparable results, considering the small sample size and the quality of the data. For both methods, the area under the (H)SROC curve was calculated and the results were similar. The SROC model does not take differences in sample size between study units into account, whereas the HSROC allows for both between and within study variation. While the SROC is easier to implement, analyse and interpret, the HSROC does have properties which might encourage the extra effort involved in the analysis.

INTRODUCTION

The ability to correctly diagnose a specific disease or infection has always been a central component in veterinary medicine. For many pathogens, the decision maker has a wide range of tools available covering the spectrum from clinical and pathological examinations, over serological assays for antibody detection, to direct identification of the pathogen. Furthermore, diagnostic tests are being used in a wider range of settings: confirmation of clinical cases, identification of subclinical infections, surveillance, certification of disease freedom, etc. The World Organisation for Animal Health (OIE) has acknowledged this and at its 71st Annual General Session in May 2003, the International Committee of the OIE adopted Resolution No. XXIX,, which introduced the 'fitness for purpose' as a criterion for validation. The 'fitness for purpose' implies that to properly evaluate a diagnostic test, the context for its application must be considered, so that the condition detected by the test reflects the purpose for which the test is intended to be used. However, the vast array of available diagnostic tools and their many purposes has further increased the need for good studies on the reliability and performance of the available tests.

Design and analysis of diagnostic test evaluation studies has become part of the standard curriculum in many epidemiological basic courses. Traditionally, the performance of a diagnostic test in an epidemiological setting is defined, conditionally on the disease status, as the sensitivity, $Se = Pr(T+|D+)$, and specificity, $Sp = Pr(T-|D-)$, where D describes the condition to

* Nils Toft, Danish Meat Association, Vinkelvej 11, Dk-8620 Kjellerup, Denmark, E-mail: nto@danishmeat.dk

be detected (the truly "diseased"), and T the test-result. In the classic design of a test evaluation, a perfect reference test, often referred to as a 'gold standard', is used to discriminate between truly diseased and truly non-diseased test individuals, which are subsequently tested with the diagnostic tests under evaluation. This approach has the advantage that an exact definition of a 'diseased' individual can be made, but the disadvantage that often there is a discrepancy between the defined disease condition and the condition relevant for the decision problem, i.e. the 'fitness for purpose' is not always met by the definition of 'disease'. For example, classical test evaluation studies often rely on agent detection methods to establish the true disease status, when serological assays are evaluated. While agent detection methods may be reliable to establish the "gold standard" for detection of infectious animals, it is usually a poor 'gold standard' for detection of infected animals or animals that are clinically affected. As a consequence, infected animals can often only be included in the study if they are infectious, i.e. shedding or excreting the agent, thereby potentially introducing selection bias. Therefore, in recent years, latent class analysis (Hui and Walter, 1980) has been widely accepted as an alternative method, where tests are evaluated in the absence of a 'gold standard'. This approach somewhat circumvent this selection bias, but at the cost of introducing a more abstract disease definition, which partly depends on the tests used in the evaluation (Toft et al., 2003).

In Greiner and Gardner (2000) it is advocated that tests essentially always should be evaluated in the population where it is intended to be used. Hence, there is an apparent need for a new evaluation of a diagnostic test whenever the purpose or target population changes. In practice, there is a cost issue and the proper evaluation of a diagnostic test is very expensive and time consuming; therefore alternatives are needed. One possibility is the use of critical reviews and potentially meta-analyses of published test evaluation studies to disseminate available information. Meta-analysis is a well established statistical discipline with a good theoretical foundation, but its practical application is often diminished somewhat by the lack of good data. The same can be said about meta-analysis of diagnostic test evaluation studies, where there are several examples of critical reviews within human medicine, but few that go beyond a critical subjective comparison. Considering the points made in the previous paragraphs, it seems that formal comparisons of Se and Sp across published studies should be discouraged and most often a critical review would exclude all but a few studies from further comparison. However, diagnostic test evaluations are improving, due to more stringent review procedures of submitted manuscripts, better reporting due to initiatives such as STARD and OIE guidelines, courses and workshops, such as the DTE-workshop series. Furthermore, even if formal inference should be avoided, it can often be justified to summarise Se and Sp across studies, just to give an idea of the potential of a diagnostic technique when applied to a specific purpose. Simple averaging of Se and Sp across studies will generally not work, since a high Se is usually achieved by lowering the requirements for Sp. Assume that three studies are reported on the exact same serological assay, differing only in their selected threshold for determining a test positive. The estimated pairs of Se and Sp are: (0.25,0.95), (0.95,0.25) and (0.75,0.75) reflecting scenarios where the ability to rule-out disease, rule-in disease or a trade-off were deemed important. The average pair of Se and Sp is (0.65,0.65), which does not summarise the properties reported by the three studies well. Clearly, there is a need for more adequate means of summarising data. One such method is the use of the summary ROC (SROC) curves which are analogous to receiver operating characteristics (ROC) curves known from evaluation of e.g. serological assays (Greiner et al.,, 2000).

The objective of this study was to present graphical and model-based approaches to estimate, summarise and compare SROC curves as a means to synthesize evidence from diagnostic test evaluation studies. To illustrate the concepts, we present the synthesis of a critical

review of accuracies of ELISA when used for detection of different stages of paratuberculosis in cattle and compare various SROC techniques based on the derived set of comparable studies.

MATERIALS AND METHODS

Critical review of ELISA for detecting paratuberculosis in cattle

Paratuberculosis is a chronical infection, which is of particular concern in ruminants. The infection, which is caused by *Mycobacterium avium* subsp. *paratuberculosis* (MAP), may develop slowly over several years. In Nielsen and Toft (2008), a critical review of accuracies of antibody ELISAs, interferon-γ assays and faecal culture techniques for ante mortem diagnosis of paratuberculosis was conducted. The control of paratuberculosis exists at many levels ranging from culling of individual clinically affected cattle to national control and eradication programmes. For the latter to be efficient, it is necessary to be able to identify MAP-free cattle (or herds), which is a different purpose than confirming a suspicion based on clinical signs. For other purposes, such as trying to reduce within herd transmission, it may be sufficient to be able to identify infectious cattle. In Nielsen and Toft (2008), three different stages of paratuberculosis were defined as target conditions for diagnostic test evaluations: 1) affected, i.e. cattle with clinical signs such as diarrhoea (persistent or intermittent), chronic weight loss or reduced milk production; 2) infectious, i.e. cattle shedding MAP at the time of testing; and 3) infected, i.e. cattle where entrance and persistence of MAP have lasted long enough to give an immune response at any time during their life. For the latter; there is no time-specific cut-off for this event to occur. It is assumed that an infection in a cow persists for life. The target condition 'infected', includes 'infectious', which again includes 'affected' cattle. For many of the reported studies, different target conditions where defined for the evaluation of Se and Sp, respectively.

For details about the review procedure including the choice of covariates and criteria for exclusion/inclusion of studies, we refer to Nielsen and Toft (2008). For the present study, it suffices to summarise the results with respect to ELISA used in cattle. Only studies, where both Se and Sp were evaluated and where sufficient details were given on the identification of the target condition(s) that were used in the evaluation. For each study, the sample sizes, target condition for Se and Sp, type of test (serum or milk) and other covariates were recorded.

Summary ROC (SROC) curves

Gatsonis and Paliwal (2006) stated that the display of paired estimates of Se and Sp in ROC-coordinates, i.e. (1-Sp,Se), is a key step in the process of statistical analysis, and that such plots ideally should include error bars for the two estimates. The SROC curve is an easy way to construct a graphical summary of such (1-Sp,Se) estimates. The first version of the simple SROC curve is attributed to Moses et al. (1993) and has been discussed in several publications (Gatsonis and Paliwal, 2006; Walter, 2002). The units of interest in any meta-analysis are the individual studies. In the simple SROC analysis, the *i*th study contributes an estimate of Se_i and Sp_i, which are transformed into the new variables S_i and D_i, defined as follows:

$$D_i = \text{logit}(\text{Se}_i) - \text{logit}(1-\text{Sp}_i) \qquad \text{and} \qquad S_i = \text{logit}(\text{Se}_i) + \text{logit}(1-\text{Sp}_i) \tag{1}$$

where D_i is equivalent to the diagnostic log-odds ratio, which is a measure of the test's accuracy in discriminating infected from non-infected. S_i can be interpreted as a measure of the diagnostic

threshold, with high values corresponding to liberal inclusion criteria for infected subjects (Walter, 2002).

The next step is to fit the linear regression model

$$D_i = a + bS_i + (c_1X_1 + c_2X_2 + \cdots + c_kX_k) + \varepsilon_i \tag{2}$$

where $(c_1X_1+c_2X_2\ldots)$ represent possible covariates X and their effects c and ε_i is the error term. The coefficient b represents the dependence of the test accuracy on the threshold. For the special case $b = 0$, the studies are homogenous and can be summarised by an overall diagnostic log-odds ratio a, which is assumed to be constant across all studies. When $b \neq 0$, the studies are heterogeneous and a represents the diagnostic log-odds when $S=0$. If $|b| > 1$ the SROC curve derived from the above model behaves counter-intuitively and such values indicate that the simple linear regression model does not fit adequately. If a suitable fit is obtained, the SROC curve can be derived by reversing the transformations in Eq. 1 and 2:

$$Se = \frac{\exp\left(\frac{a}{1-b}\right)\left(\frac{1-Sp}{Sp}\right)^{(1+b)/(1-b)}}{1+\exp\left(\frac{a}{1-b}\right)\left(\frac{1-Sp}{Sp}\right)^{(1+b)/(1-b)}} \tag{3}$$

Using Eq. 3, Se can be calculated for every value of (1-Sp) and the SROC curve can be plotted. The covariates have been ignored in Eq. 3 and subsequent formulas, but their inclusion are straightforward. For example in Eq. 3 covariates are added by replacing a by $a+(c_1X_1+c_2X_2\ldots)$. Alternatively, the covariates are added to the intercept in the output from the linear regression model, depending on the capabilities of the software chosen for modelling.

As in the ordinary ROC analyses, there are overall summary measures of the curve's behaviour, which may be relevant. One of the natural measures is the area under the curve (AUC), which can be calculated by integrating Eq. 3 over the possible range for (1-Sp), i.e.:

$$\mathrm{AUC} = \int_0^1 \frac{\exp\left(\frac{a}{1-b}\right)\left(\frac{x}{1-x}\right)^{(1+b)/(1-b)}}{1+\exp\left(\frac{a}{1-b}\right)\left(\frac{x}{1-x}\right)^{(1+b)/(1-b)}}\,\mathrm{d}\,x \tag{4}$$

Although there is no analytical solution to this integral, it is straightforward to calculate using numerical integration using e.g. the R package (R Development Core Team, 2007). The AUC represents the average value of Se over all possible (1-Sp) values, and is usually referred to as the probability that the test correctly ranks a random pair of infected and non-infected subjects based on their observed test values. The AUC is widely used in traditional ROC analyses, but somewhat more concern has been raised when used in SROC analyses, mostly due to the potential lack of cover of the full range of (1-Sp) values.

The SROC model is simple and adequate for summarising paired estimates of Se and (1-Sp) across studies. Often the preceding critical review process will have identified several critical issues regarding the quality of the test evaluation studies. This will discourage from more sophisticated analyses, since the graphical summary provided by the SROC curve already should be interpreted with caution if the quality of the data is poor. However, the SROC presented here does have some important limitations. Most importantly, the model is unable to distinguish between within-study and between-study variability, thus giving equal weight to all pairs of (Se,1-Sp) despite potentially large differences between studies with respect to sample sizes. While the SROC model can be fitted using weighted analysis, it is generally discouraged (Walter, 2002) as it produces biased estimates. Therefore, a hierarchical SROC model (HSROC) would be preferred.

The HSROC model can be regarded as having two levels corresponding to within and between study variability (Rutter and Gatsonis, 2001). Here we adopt the formulation of the HSROC from Macaskill (2004): for the *i*th study, the number of test positives for the infected (*I*) and non-infected (*N*) test subjects, $t_{ij:}$ *j=I,N,* respectively, are assumed to follow binomial distributions

$$\begin{aligned} t_{ij} &\sim \text{Bin}(p_{ij}, n_{ij}) \; j = I, N; i = 1,2,... \\ \text{logit}(p_{ij}) &= (\theta_i + \alpha_i d_{ij}) \exp(-\beta d_{\text{ij}}) \\ \theta_i &\sim \text{N}(\mu_\theta, \sigma_\theta^2) \\ \alpha_i &\sim \text{N}(\mu_\alpha, \sigma_\alpha^2) \end{aligned} \tag{5}$$

where p_{ij} is the probability of a positive test result for the *i*th study and *j*th infection status; n_{ij} is the sample size for the *i*th study and *j*th infection status; d_{ij} is the true disease status for the *i*th study and *j*th infection status (coded as 0.5 for *j=I* and -0.5 for *j=N*); and the random effects (θ_i and α_i) are assumed to be independent and normally distributed. Thus each study has its own implicit threshold θ_i, estimating the average log odds of a positive test result for the infected and non-infected groups, and diagnostic accuracy α_i, estimating the expected diagnostic log odds ratio. The scale parameter β allows the accuracy to vary with implicit threshold, thereby allowing asymmetry in the SROC. The scale must be held fixed since each study only contributes one point to the SROC curve. Hence, the association between threshold and accuracy must be derived from the studies considered jointly. From Eq. 5 the HSROC curve can be derived as:

$$\text{Se} = \frac{1}{(1 + \exp(-(\mu_\alpha \exp(-0.5\beta) + \text{logit}(1 - \text{Sp}) \exp(-\beta))))} \tag{6}$$

by varying (1-Sp) across the relevant range (e.g. the interval [0,1] for the full curve). For the HSROC curve, the AUC can be calculated by integrating Eq. 6 similar to the SROC model (Eq. 3 and 4).

All statistical analyses and data management was performed using R (R Development Core Team, 2007) except the estimation of the HSROC parameters, which was carried out using PROC NLMIXED in the SAS software, Version 9.13 of the SAS System for Windows XP.

RESULTS

The review process and subsequent exclusion of studies which did not meet the requirements resulted in a dataset with 36 test evaluations at least partially suitable for further analyses. Of the 36 evaluations, 31 were concerning serum antibody ELISAs and 5 were milk antibody ELISAs. Four studies used the target condition 'infected' (E) for both the evaluation of Se and Sp, 19 used the target condition 'infectious' (I) for both evaluations and 13 used 'infectious' for the Se evaluation and 'infected' for the Sp evaluation. We will refer to these three groups as EE, II and IE, respectively, throughout the rest of this paper. Sample sizes varied substantially between studies, the number of subjects with the target condition present ranged from 8 to 415 (median sample size 114) and subjects without the target condition ranged from 16 to 1751 (median sample size 346). As suggested by Gatsonis and Paliwal (2006), the starting point of the statistical analysis is a forest plot (Fig. 1) giving the Se and Sp with error bars (showing the 95% confidence interval, using the traditional normal approximation formula).

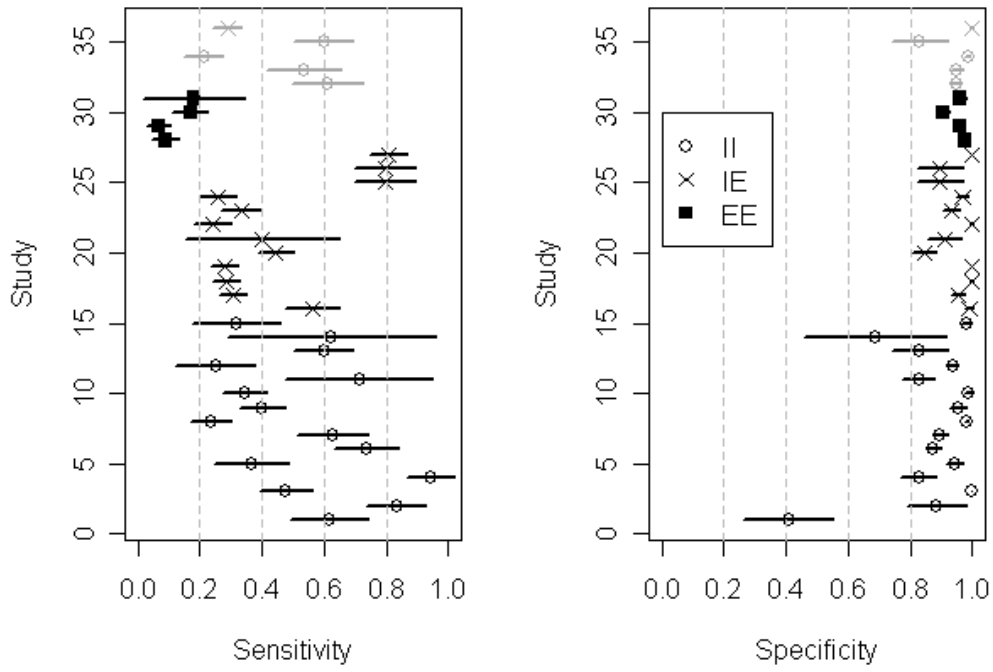


Fig 1. Forrest plots presenting the Se and Sp of the 36 studies used for illustration of the different summary ROC techniques. Different symbols are used to distinguish between the three classes of target conditions (II, EE and IE). Black colour represents serum ELISA and grey represents milk ELISA. The bars give the width of the 95% confidence interval, symbols the mean estimate.

For the SROC analysis, the initial model included test-type (milk/serum) and target conditions as covariates including possible second-order interactions. As suggested by Walter (2002), 0.5 was added to each cell in the underlying 2x2 tables for each study to avoid problems with zero-cells before calculating Se and Sp (i.e. to avoid problems occurring by Se or Sp equal to 1). Only an effect of the target condition EE on the diagnostic accuracy remained after reducing the model by removing (the most) non-significant effects, one at a time (Table 1). In the presentation of the results, the effect of the target condition was incorporated in the estimate for the intercept to give an intercept for target condition EE and II+IE, respectively. The estimate of the slope (b=-0.32) was non-zero, suggesting heterogeneity between studies. For the HSROC the initial model included all main effects but no interactions. The model was reduced in the same manner as the SROC, except that variance parameters in the random effects were tested using Akaike's Information Criteria, since estimated P-values for such parameters should be interpreted with caution. The scale parameter (β) was removed from the model second to last, and subsequently tested in the final model. However, it was not found to be significant, thus in the HSROC model the included studies were found to be homogenous (Table 1). Based on the estimated parameters for the two models, the corresponding SROC and HSROC curves were plotted (Fig. 2) along with the estimated Se and (1-Sp) for the studies.

Table 1. Estimates of the SROC and HSROC parameters and AUC for the final models using 36 ELISA evaluations.

ANALYSIS	PARAMETER	ESTIMATE	S.E.	P-VALUE	AUC
SROC	$a_{II,IE}$	2.10	0.36	<.0001	0.80[a]
	a_{EE}	-0.51	0.77	0.0004	0.42[b]
	b	-0.32	0.10	0.0020	
HSROC	μ_θ	-1.63	0.18	<.0001	0.87
	σ_θ^2	1.08	0.28	0.0004	
	μ_α	2.82	0.23	<.0001	
	σ_α^2	1.66	0.48	0.0016	

[a] AUC for SROC curve corresponding to the II and IE target conditions;
[b] AUC for SROC curve corresponding to EE target condition.

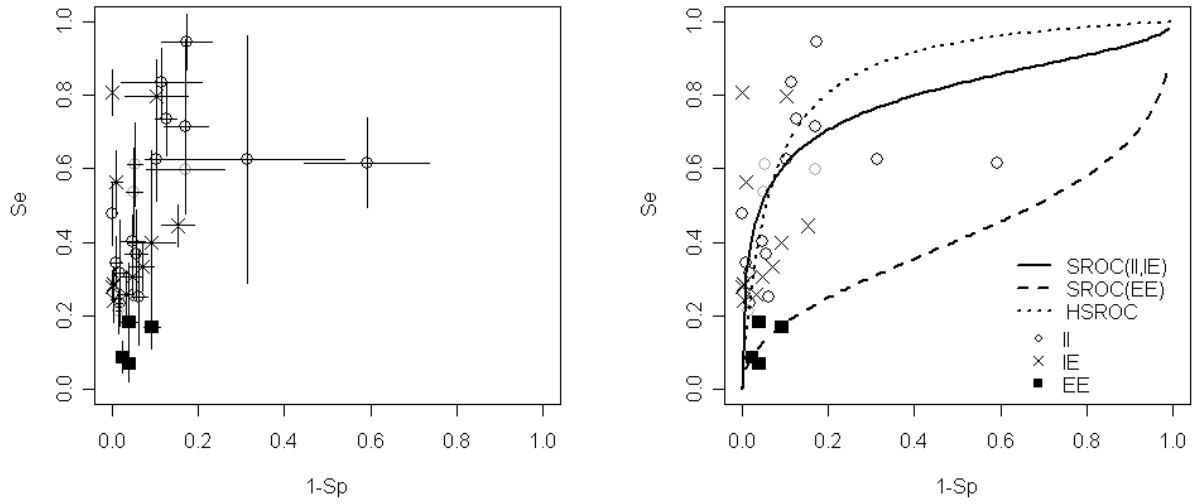


Fig. 2. The SROC and HSROC curves and original data points. II, IE and EE refer to the three different target conditions used in the analyses. In the left plot, 95% confidence intervals for the estimates are shown.

DISCUSSION

In this study we have compared simple SROC and hierarchical SROC models on a dataset of 36 diagnostic test evaluation studies for antibody ELISAs for ante mortem diagnosis of paratuberculosis. Three different definitions of target condition were used in the evaluation study. The two models gave somewhat different, although comparable results.

Using the simple SROC model based on ordinary linear regression, we found that there was a difference between the diagnostic log odds ratio for different test evaluation studies when the target condition for the evaluation of Se differed. For studies where the target condition for Se was 'infectious' the diagnostic log odds ratio as well as the AUC = 0.80 suggested reasonable discriminatory power, whereas the diagnostic log odds ratio and AUC = 0.42 for evaluation studies using the target condition 'infected'. This indicates that this target condition does not apply well to antibody ELISAs. The estimated slope (*b*) was significantly different from zero, which indicated heterogeneity between studies, implying that the diagnostic log odds ratios differed between studies. The estimated AUC of 0.42 for the target condition EE group should only be considered with extreme caution. Only 4 studies were used the EE classification and they all were located at the lower left corner of the SROC plot, leaving most of the drawn SROC curve unsupported. This is one of the major concerns when using and interpreting AUC in meta-analysis of diagnostic test evaluation data. Some studies have suggested the use of the partial area under the curve, i.e. only the fraction of the SROC which is supported by the data. In Walter (2005) the partial AUC is addressed in detail. The conclusion being that although it does have some appealing properties, the full AUC should be used predominantly. For our study, we might consider to ignore the results, since they are based on only 4 studies. It is notable, however, that these studies have remarkable similar features: a poor Se and a no particularly impressive Sp.

In the HSROC analysis, no differences were found between target conditions and there was no evidence of heterogeneity, i.e. the scale parameter β was not significantly different from zero. This implies that the HSROC curve is symmetric and the diagnostic log odds ratio can be summarised by μ_α. The AUC was estimated to be slightly higher than for the SROC analysis and the curve increased more steeply towards the top. The sample sizes differed between studies, as seen by the variation in the width of the error bars in Fig. 1 and Fig. 2. Therefore, the multi-level model allowing for between and within study variation appeared to be more appealing despite its somewhat more complicated implementation. The benefit of accounting for differences between sample sizes is easily seen in Fig. 2, where the SROC for the IE and II target conditions is clearly influenced more by the two rightmost data point than the HSROC. In the SROC analysis these two points have the same influence as data from any other study, since only Se and Sp estimates are used. In the HSROC analysis, the variability due to the relatively small sample sizes is taken into account. The result is the visible difference in the shape of the curves at the top-right part and the improved AUC of the HSROC compared to the SROC. The differences in the final models with respect to included covariates and shape parameter is probably to some degree the result of the relatively small number of studies included in the analysis. From a logical point of view, 'target condition' should be a significant covariate. The HSROC model with effect of target condition EE and a shape parameter gave estimates close to those obtained in the SROC model, but the effects were not found to be statistically significant.

The major problem in this analysis is the same as in almost all meta-analyses: the quality of the design and reporting of the individual test evaluation studies makes it very difficult to justify using more than a very small subset for further inference, which then in practice prevents any analysis due to the lack of data. The inclusion criteria which were presented in the review of Nielsen and Toft (2008) included 6 different requirements about the disease definition (target condition), study population, sampling protocol etc. They represented what was seen as the minimal possible set of selection criteria. Using these it was possible to include a reasonable number of studies for further analysis, rather than excluding almost all of 153 diagnostic test evaluation studies.. Although the latter might have been scientifically sounder, it could also be considered somewhat counter-productive. It was chosen to favour 'quantity' by including studies where the quality of reporting was somewhat less than desirable, rather than 'quality' by only including studies where the detail of reporting made a more rigorous judgement of quality possible. As a consequence, most of the included studies lacked reliable information about potentially relevant covariates such as the specific antigen used in the test, the age-distribution of the test subjects, etc.

Despite the problems with data quality, summary ROC analyses are useful, as they help to provide graphical summaries of the available data, thereby helping the investigators to get an overview of the possibilities with the different techniques. The SROC curve is an indicator of the potential of a diagnostic technique. This should be taken into account when designing, e.g. herd certification schemes based on ELISAs. It would be unreasonable to assume that a local in-house ELISA can be perfected beyond what is suggested possible by the published studies. As properly conducted test evaluation studies require large sample sizes of randomly selected naturally infected (and randomly selected 'naturally' non-infected) test subjects, they are costly and time consuming. Furthermore, paired comparisons are being increasingly advocated as a the way forward, thereby necessitating multiple tests being simultaneously evaluated, hence further adding to the cost of such evaluations. As a consequence, one is forced to make the most of available data and meta-analysis is one way to achieve this.

The overall poor quality of data might suggest that the simpler SROC analysis is to be preferred over the more complicated HSROC. Indeed the SROC has some very appealing properties. Its simplicity allows it to be implemented in almost any available statistical software including Excel spreadsheets. The linear regression analysis is basic statistics and is familiar to most investigators. However, the HSROC does allow for the random variation between studies (which is inherently more realistic with the observed lack of common standards) and does account for differences in sample sizes which we have demonstrated to have at least a visible effect on the results. Furthermore, the HSROC can be augmented to account for the correlation between tests in a paired design, when two or more tests are evaluated on the same dataset, by imposing a suitable correlation structure within the individual studies. Similarly, multiple evaluations of the same test, e.g. at different cut-off points, can be included by adding a suitable correlation structure.

It is possible to calculate the standard errors of the AUCs and thus make a formal comparison of these summaries. The calculations are rather tedious (see e.g. Walter (2002)) and requires the use of numeric integration combined with the estimated standard errors of the parameters in the SROC or HSROC models. Alternatively, one might obtain Monte Carlo estimates of the standard errors by simulation.

We have used classical statistical methods for the formulation and analysis of the models in this study. Naturally, Bayesian methods are realistic alternatives and such models are discussed in e.g. Hellmich et al. (1999), Macaskill (2004) and Rutter and Gatsonis (2001). The models are straightforward to implement in e.g. OpenBugs (Thomas et al., 2006). We have refrained from doing so in this study in order to avoid the issue of obtaining prior information. The HSROC can be implemented in R as well as SAS, and PROC NLMIXED was only chosen because the authors were more familiar with this software and sample code for SAS was included as an appendix in Macaskill (2004).

To conclude, the SROC and HSROC are useful tools for presenting graphical summaries of available data in reviews of diagnostic test evaluation studies. Combined with the use of colour and symbols in the drawings, it helps visualizing the possibilities for the different techniques.

ACKNOWLEDGEMENTS

This study was co-funded by the European Commission within the Sixth Framework Programme, as part of the project ParaTBTools (contract no. 023106 (FOOD)).

REFERENCES

Gatsonis, C. and Paliwal, P. (2006). Meta-analysis of diagnostic and screening test accuracy evaluations: methodologic primer. American Journal of Roentgenology. 187, 271-281

Greiner, M. and Gardner, I.A. (2000). Epidemiologic issues in the validation of veterinary diagnostic tests. Preventive Veterinary Medicine 45, 3-22

Greiner, M., Pfeiffer, D. and Smith, R.D. (2000). Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. Preventive Veterinary Medicine 45, 23-41

Hellmich, M., Abrams, K.R. and Sutton, A.J. (1999). Bayesian approaches to meta-analysis of ROC curves. Medical Decision Making 19, 252-264

Hui, S.L. and Walter S.D. (1980). Estimating the error rates of diagnostic tests. Biometrics 36, 167-171.

Nielsen, S.S and Toft, N. (2008). Ante mortem diagnosis of paratuberculosis: A review of accuracies of ELISA, interferon-γ assay and faecal culture techniques. Veterinary Microbiology, (In press) doi: 10.1016/j.vetmic.2007.12.011

Macaskill, P. (2004). Empirical Bayes estimates generated in a hierarchical summary ROC analysis agreed closely with those of a full Bayesian analysis. Journal of Clinical Epidemiology. 57, 925-932

Moses, L.E., Littenberg, B., and Shapiro, D. (1993). Combining independent studies of a diagnostic test into a summary ROC curve: data analytical approaches and some additional considerations . Statistics in Medicine 12, 1293-1316

R Development Core Team. (2007). 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

Rutter, C.M. and Gatsonis, C.A. (2001). A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. Statistics in Medicine. 20, 2865-2884

Thomas, A., O'Hara, B., Ligges, U. and Sturtz, S. (2006) Making BUGS Open. R News 6: 12-17

Toft, N, Nielsen, L.R. & Højsgaard, S. (2003). Exploring disease definitions of Salmonella Dublin in latent class analysis. In: Proceedings of the 10th International Symposium on Veterinary Epidemiology and Economics, Vina del Mar, Chile. November 17 – 21 – 2003

Walter, S.D. (2005). The partial area under the summary ROC curve. Statistics in Medicine. 24, 2025-2040

Walter, S.D. (2002). Properties of the summary receiver operating characteristic (SROC) curve for diagnostic test data. Statistics in Medicine. 21, 1237-1256

SURVEILLANCE

USING SCENARIO TREE MODELLING FOR THE EVALUATION OF SURVEILLANCE STRATEGIES FOR ZOONOSES IN DISEASE-FREE AND ENDEMIC SITUATIONS

D.C. HADORN*, J. ZINSSTAG, S. SERIC HARACIC, AND K.D.C. STÄRK

SUMMARY

In recent years, the surveillance of animal diseases has been influenced by the threat of emerging infectious diseases of animal origin. Zoonotic diseases account for the majority of all emerging infectious diseases. Therefore, veterinary services are more and more challenged with public health issues. This stands in contrast with the increasingly limited resources for disease surveillance. Bovine tuberculosis (bTB) and *Brucella melitensis* (Bm) are two important zoonotic diseases occurring worldwide and causing serious public health and economic damage. The surveillance of such zoonoses is complex and resource consuming. In order to efficiently economize available resources, it is important to assess and evaluate the detection performance of each surveillance system component and the overall performance of the surveillance system. In this paper, an approach using scenario tree modelling is described for the example of bTB surveillance in Switzerland and Tanzania and the surveillance of Bm in Switzerland as well as Bosnia and Herzegovina.

INTRODUCTION

Due to the global network of trade with animals and animal products and due to climate change, emerging infectious diseases are increasingly present in both developed and developing countries. Most emerging infectious diseases are zoonoses affecting both human and animal populations (Slingenbergh et al., 2004). Therefore, veterinary services are more and more challenged with the surveillance of zoonoses and consequently with public health issues. Because of such increasing tasks but decreasing resources, new approaches for the evaluation and optimization of zoonotic disease surveillance systems need to be developed. According to the World Animal Health Organization (OIE), a *surveillance system* for an infectious animal disease is defined as a method of surveillance that may include one or more component activities that generates information on the health, disease or zoonosis status of animal populations (Terrestrial Animal Health Code of the OIE, 2007). The approach described in this paper is based on the analysis and evaluation of all surveillance system components (SSCs) of one surveillance system. Analyzing and evaluating the individual SSCs using the scenario tree methodology, the whole surveillance system can be evaluated.

In general, SSCs may be based on two different surveillance approaches, i.e. 'active' or 'passive' surveillance: Lilienfeld & Stolley (1994) describe *active surveillance* as the regular periodic collection of samples or case reports from veterinary health authorities. Because this

* Daniela C. Hadorn, Swiss Federal Veterinary Office, Schwarzenburgstrasse 155, Switzerland - 3003 Bern, Switzerland. Email: daniela.hadorn@bvet.admin.ch

pproach is very cost-intensive, especially for rare health-related events where a large sample size is required, the application of *risk-based surveillance* is a promising approach. In this case, sampling may be targeted on high-risk populations where the probability is highest to find the disease given that the disease is present (Stärk et al., 2006). *Passive surveillance* is described as the reporting of clinical suspect cases to the health authorities (Lilienfeld & Stolley, 1994). Because in this case, animals or samples are only tested if a disease suspicion exists, this approach is cost-saving and therefore attractive to be integrated in a surveillance system. But the detection performance of a passive SSC is influenced by the probability of infected animals to show clinical symptoms, by the disease awareness of persons responsible for reporting and by the sensitivity of the applied diagnostic test. Therefore, it is difficult to estimate the detection performance of a passive SSC. Another disadvantage of this approach is under-reporting. Therefore, in most of the cases, passive surveillance is not suitable to be the only surveillance activity in order to provide reliable information on the actual disease status of a population.

In this paper, an approach is described for the quantification and evaluation of a surveillance system for two important zoonoses, namely bovine tuberculosis (bTB) and *Brucella melitensis* (Bm). The situation is analyzed in three different countries comparing the situation of disease freedom and endemic situations. Switzerland (Switzerland) serves as the example for disease-free situation for both bTB and Bm. Tanzania provides the example of an endemic situation of bTB in a low-income country, and Bosnia and Herzegovina (BH) is the example for a Bm-endemic country situation.

MATERIALS AND METHODS

Scenario tree modelling

Because passive surveillance is an ongoing and cost-saving process and therefore an interesting tool to save resources, veterinary services should be able to assess the contribution of passive SSCs to the overall performance of a surveillance system. If the detection performance of passive SSCs like clinical surveillance is already very high due to a high probability for detectable clinical symptoms in infected animals and high disease awareness in farmers and veterinarians, there may be no need for a cost-intensive, active survey. But if passive SSCs are not very sensitive, it may be reasonable to conduct additional surveys. The sensitivity of an SSC expresses the probability that at least one unit would be detected through the corresponding SSC given that the disease in the country or region is equal to or above the design prevalence. In this paper, detection performance is used similar to the sensitivity of an SSC. An important input parameter for passive SSCs is the disease awareness of all involved persons. Because disease awareness is influenced by many factors like knowledge of farmers and veterinarians about a disease, by management factors like compensation payment for culled animals and economical drawbacks from reporting suspect cases, it is difficult to quantify disease awareness. Within the scope of this project, four levels of disease awareness were arbitrarily defined and used (Table 1). In general, the quantification and evaluation of disease awareness needs more consideration and research in the future.

Table 1. Qualitative description of the disease awareness categories used in this project.

CATEGORY	DISTRIBUTION	LEVEL
Low disease awareness	*RiskPert* [0.1; 0.2; 0.3]	L
Medium disease awareness	*RiskPert* [0.4; 0.5; 0.6]	M
High disease awareness	*RiskPert* [0.7; 0.8; 0.9]	H
Very high disease awareness	Fixed value [1.0]	VH

Passive SSCs depend not only on disease awareness of persons responsible for reporting, but also on other factors like probability for infected animals to show clinical symptoms. Therefore, stochastic scenario tree modelling offers a valuable approach to identify and depict each step in the process from disease presence to disease detection and to calculate the probability to detect at least one case with a certain SSC. A pivotal point of the scenario tree approach of Martin et al. (2007) is that risk factors for an infection with a specific disease agent are taken into account. Therefore, the animal population under survey can be stratified according to its risk of infection. Additionally, the impact of population strata on the disease situation of the country or region is incorporated by implementing relative risks (RR) for different sub-units of the population. Hadorn & Stärk (submitted) showed that the stochastic simulation model can also be used to compare the performance of random surveys (RS) and targeted sampling (TS) and to compare and analyze different sampling designs with regard to cost-efficiency. In this work, Microsoft *Excel* and Palisade *@RISK* (www.palisade.com) were used as the modelling software and simulations were run with 10,000 iterations each.

The first step in analyzing a surveillance system using stochastic scenario tree modelling is to collect information on the disease and disease situation in the corresponding country or region as well as to identify the risk factors for infections in the animal population which has to be surveyed. The second step includes the identification of the SSCs that contribute information on the disease status of the population. All SSCs are analyzed step by step using the scenario tree approach, the required input parameters are collected and the detection performance of each SSC calculated.

Surveillance system for bovine tuberculosis in Switzerland and in Tanzania

Switzerland: Switzerland is officially free from bTB, so the design prevalence is set to be 0.2% according to the OIE threshold prevalence for disease free countries. The following three possible pathways exist for bTB incursion in Switzerland (all relative risks RR given in brackets are based on expert opinion): The risk of disease introduction through infected wildlife from abroad ($RR_{WILDLIFE}$ = 5) or through trade with infected cattle (RR_{TRADE} = 10) and the risk of disease transmission from bTB-infected humans to cattle (RR_{HUMAN} = 1.5). Furthermore, the surveillance system for bTB in Switzerland consists of passive clinical surveillance on farm (CLIN) and at the slaughterhouse (SLI) as well as human clinical case surveillance (Human Sentinel Surveillance HSS) with tracing-back on suspected cattle farms. The disease awareness levels in Switzerland for CLIN are assumed to be medium and low for farmers and veterinarians, respectively. Disease awareness is estimated to be medium in meat inspectors for SLI and low for the cooperation between human and animal health authorities in HSS. There are currently no active SSCs conducted in Switzerland. But a hypothetical random sampling (RS) with 1,500 herds out of the general population was compared to a theoretical targeted sampling

(TS) with 500 and 50 cattle herds out of the high-risk group 'WILDLIFE' and a more targeted high-risk group 'WILDLIFE & TRADE', respectively.

Tanzania: bTB is largely present in most of the African cattle and wildlife populations. The importance of this zoonosis in Africa is assumed to be increasing especially due to the increasing HIV infections in humans (Kazwala et al., 2001). There seem to exist various potential risk factors for bTB infection in Tanzanian cattle like breed (exotic breed seem to be more susceptible to bTB than indigenous zebu cattle), cattle management systems, contact with wildlife and geographical set-up. But currently, data are not sufficient to identify distinct high-risk population strata and their corresponding RR. Comparing the surveillance system of Switzerland with Tanzanian situation, it can be stated that there is no effective CLIN with regard to bTB, especially in remote areas, because of lack of veterinary service. Furthermore, bTB symptoms like loss of weight and coughing are unspecific, and there exist a lot of tropical diseases causing similar symptoms which are more known than bTB (Pers. comm. B. Bonfoh, 2007). SLI is applied in Tanzania but at an unknown and probably limited extent because most animals are slaughtered in backyards without meat inspection. HSS may be helpful for bTB surveillance in Tanzania (expert opinion). But it has to kept in mind that risk factors for human infection with *M. bovis* are related to poverty and that there exist a lot of infection sources other than the own cattle (purchase of milk products on markets, consumption of raw milk and meat in other households). This impedes a direct and straightforward human-cattle-surveillance. Concerning RS and TS, no national surveillance program exists so far. But the authors propose that risk-based surveillance should be considered by Tanzanian veterinary service to generate reliable information on the bTB status of Tanzanian cattle population.

Surveillance system for *Brucella melitensis* in Switzerland and in Bosnia and Herzegovina

Generally speaking, sheep and goat flocks belong to the same epidemiological unit for the surveillance of Bm in these countries. Therefore, sheep and goat flocks are not treated separately in the scenario trees and the term 'flock' includes sheep and goat species.

Switzerland: Switzerland is free from Bm and therefore, the design prevalence is set to be 0.2% (according to the OIE threshold prevalence for disease free countries). The only identified risk factor for disease introduction is the import of potentially infected sheep and goats, respectively. The uncertainty with regard to the RR of the high-risk group with imported small ruminants "IMP" was taken into account using a RiskPert distribution (RR_{IMP}= *RiskPert* [2.5; 10.4; 43]) (Lithg-Pereira et al., 2004). The surveillance system for Bm in Switzerland consists of a passive SSC, namely abortion testing ABT, and an active component with annual RS. The assumptions concerning the disease awareness in Swiss farmers and veterinarians for ABT are low and medium, respectively. The sample size used for RS was assumed to be 1,500 flocks. A hypothetical TS with a 10 fold smaller sample size of 150 flocks was also considered to analyze the potential benefit of sampling in high-risk population strata "IMP" compared to RS in the general small ruminant population.

Bosnia and Herzegovina: Bm is considered to be endemic in BH since the prevalence in the national small ruminant population is > 0.2%. But for the purpose of analysis and comparison of the various SSCs, the surveillance system of BH was analyzed for a design prevalence of 0.2%. The endemic Bm situation in BH has an influence on the disease awareness of farmers and veterinarians which is obviously higher in this country than in Switzerland. Namely, the disease awareness in farmers, depending on the farm management system, is assumed to be low-medium in the fenced range farm management system (transhumance) and medium in the conventional

management system. Veterinarians are assumed to have a medium-high disease awareness in BH. The identified risk factors for Bm infection in BH are fenced range farm management (FRFmng) with an assumed RR of *RiskPert* [5.0; 6.7; 22.6] (Pers. opinion S. Seric, Kabagambe 2001; Al-Talafhah 2003) and flocks with a previous brucellosis history (Bhis) (RR_{Bhis}= *RiskPert* [3.0; 3.4; 4.0]) (Pers. opinion S. Seric, Lithg-Pereira et al., 2001). The surveillance system for Bm in BH is composed of ABT and RS, equally to Switzerland. But as a difference to Switzerland, there are two additional passive SSCs in BH, namely tracing back of brucellosis cases in cattle (Cattle sentinel surveillance CSS) and in humans (HSS) to the potential source in small ruminants.

All SSCs listed above for the different countries were analyzed and evaluated using scenario tree modelling. The evidence provided by all passive SSCs was analyzed over one year. The detailed input parameters and the structure of the scenario trees are not shown within the scope of this paper but can be provided by the authors upon request.

RESULTS

Analysis of the bovine tuberculosis surveillance system in Switzerland and in Tanzania

Switzerland: CLIN does not provide a useful detection performance due to the lack of pathognomonical clinical symptoms for bTB infected cattle. Also HSS is a negligible information source for Swiss situation due to the very low incidence of human bTB cases (0.036/100,000). The probability to detect at least one infected bTB case in Swiss slaughterhouses at a prevalence of 0.2% is 55.6%. The detection performance of a RS with 1,500 herds is 92.4%. If TS in 'WILDLIFE' risk strata of the cattle population is conducted with a 3-fold smaller sample size than in RS, the detection performance is nevertheless higher with 96.7%. An even more targeted sampling in 'WIDLILFE & TRADE' risk strata shows that the detection performance of such a sampling design is almost as high as TS in 'WILDLIFE' but with 10% of the sample size of TS 'WIDLILFE & TRADE' (Table 2). Due to the high costs for tuberculin testing, the most cost-efficient SSC seems to be SLI. In order to improve the existing surveillance system, it is proposed to increase disease awareness in meat inspectors through information campaigns.

Table 2. Simulation results of the probability of case detection by individual surveillance system components for bovine tuberculosis surveillance in Switzerland

SURVEILLANCE SYSTEM COMPONENT	P(DETECTION)		
	5%	50%	95%
Clinical surveillance CLIN	0.002	0.006	0.010
Slaughterhouse surveillance SLI	0.397	0.556	0.727
Human sentinel surveillance HSS		<< 1	
Random sampling RS (1,500 herds)	0.911	0.924	0.936
Targeted sampling TS in risk strata 'WILDLIFE' (500 herds)	0.959	0.967	0.974
Targeted sampling TS in risk strata 'WILDLIFE & TRADE' (50 herds)	0.948	0.958	0.966

Tanzania: Discussing with experts, it becomes very clear that the detection performance of CLIN is negligible for current Tanzanian situation. Farmers and also veterinarians are not sensitized enough for this zoonosis, and no tuberculin testing is used in the field for suspect cases. Concerning HSS, there is no typing of human TB strains so far to differentiate between *M. tuberculosis*, *M. bovis* and other mycobacteria. Therefore, HSS does currently not exist in Tanzanian situation. The third passive SSC for bTB surveillance, SLI, is conducted in Tanzania to a certain percentage. But at this time, cattle population and slaughter data are not sufficient to estimate the sensitivity of SLI for Tanzania. Because there is not enough data available on the population structure and the distribution concerning the identified risk factors of Tanzanian cattle, RS and TS could not be calculated as well. In order to design a cost-efficient and significant surveillance system for bTB in Tanzania, it is proposed to collect further detailed information about risk factors for bTB infection, cattle population structure, slaughter facilities and the feasibility of HSS using Delphi panel method (Oranga & Nordberg, 1993). The second step would be to evaluate and analyze possible SSCs and to identify the most cost-efficient surveillance system.

Analysis of the *Brucella melitensis* surveillance system in Switzerland and in Bosnia and Herzegovina

Switzerland: ABT provides an assumed detection performance of 35.8% for the current estimated disease awareness in farmers and veterinarians. Testing 1,500 randomly selected small ruminant flocks results in a detection performance of 94.3%. This stands in contrast to an even higher detection performance of 96.7% for a TS of just 150 flocks selected out of the high-risk group 'IMP' (Table 3). Based on these results, the following surveillance system can be proposed to be the most cost-efficient one: to conduct an annual TS in the high-risk group 'IMP' – instead of RS – and to continue the ABT as the ongoing and basis surveillance tool for the general small ruminant population in Switzerland.

Bosnia and Herzegovina: The passive SSC of ABT reaches a detection performance of 49.3% in BH and, compared to ABT in Switzerland, is 2.7-times more sensitive than the Swiss ABT-SSC (comparing the detection performance of one single unit, data not shown). This higher probability of case detection is associated with the higher disease awareness in farmers and veterinarians. The case detection by CSS and HSS was negligible and is therefore not further discussed. The similarity in the detection performances of RS in BH compared to the one in Switzerland shown in Table 3 is due to the similar sampling protocol between BH and Switzerland. Because TS is a good methodology for early detection of disease incursion in a disease free country or region but makes no sense for the purpose of prevalence monitoring in endemic situation, no TS was calculated for BH. But in the case of an eradication strategy, it would be proposed to replace the current RS with a sampling strategy taking into account high-risk population strata.

Table 3. Simulation results of the probability of case detection by individual surveillance system components for *Brucella melitensis* surveillance in Switzerland (Switzerland) and in Bosnia and Herzegovina (BH)

COUNTRY	SURVEILLANCE SYSTEM COMPONENT	P(DETECTION)		
		5%	50%	95%
Switzerland	Abortion testing ABT	0.259	0.358	0.451
	Random sampling RS (1,500 flocks)	0.942	0.943	0.950
	Targeted sampling TS (150 flocks)	0.736	0.967	0.998
BH	Abortion testing ABT	0.354	0.493	0.613
	Random sampling RS (1,500 flocks)	0.909	0.949	0.950
	Cattle sentinel surveillance CSS		<< 1	
	Human sentinel surveillance HSS		<< 1	

DISCUSSION

The surveillance of zoonoses poses a special challenge to veterinary services. Disease can be transmitted between animals and humans and therefore, the epidemiological situation is more complex than for a disease restricted to animals. Additionally, veterinary services are not only confronted with animal disease situations but also with public health issues. Due to limited resources, all possible surveillance activities should be considered and evaluated with regard to their detection performance and the most cost-efficient surveillance strategies forced. The scenario tree modelling used in this project offers a potential approach of estimating the quality and usefulness of several different surveillance activities. The various SSCs can be assessed objectively and can also be compared between different countries. It could also be demonstrated that the evaluation of different SSCs using scenario tree modelling is useful either for disease free situation as well as for endemic situation.

Analyzing the surveillance system for bTB, it can be stated that CLIN not a useful tool for Switzerland or Tanzania. Lack of pathognomonical clinical symptoms for bTB infected cattle and the currently neglected relevance of this disease in Tanzania impede a high detection performance of this SSC. HSS is also a tool that provides a negligible probability of case detection in Switzerland and is not present at all in Tanzania. But it can be stated that the surveillance of bTB in human cases will become an important part of the general bTB surveillance in African countries. SLI is the ongoing and most important passive SSC with regard to bTB surveillance. But the sensitivity of this tool depends a lot on the disease awareness of meat inspectors and – with regard to African situation – on the percentage of animals slaughtered in slaughterhouses under the control of meat inspectors.

Comparing RS in the general Swiss cattle population with TS in two different categories of high-risk populations, it can be shown that TS is much more cost-efficient than RS. But it is important to state that the benefit of TS compared to RS depends primarily on the identified risk factors and their RRs and on the distribution of the population within the high-risk strata. If no clear high-risk strata can be identified like in Tanzania, a TS is not feasible. But as soon as there are more data available, TS could also be a cost-saving and efficient alternative for a future nationwide active surveillance strategy in Tanzania.

The surveillance systems of Bm in Switzerland and BH show similar general results. ABT as the most important passive SSC is a cost-saving and ongoing surveillance activity depending a lot on the disease awareness of farmers and veterinarians. It can be shown that ABT in BH provides a higher detection performance than ABT in Switzerland due to the endemic situation and the subsequent sensitization of people in BH. Again, the comparison of RS and TS in Switzerland shows that TS could be a cost-efficient option with the purpose of early detection of disease incursion.

Analyzing a surveillance system with the scenario tree approach also offers the opportunity of simulating and analyzing the effect of varying input parameters on the sensitivity of a surveillance system. The most influential parameters can be identified and – linked to the costs – an optimal surveillance system with regard to cost-effectiveness can be designed.

ACKNOWLEDGEMENTS

The authors thank Andy Mitchell and Martin Vordermeier from Veterinary Laboratories Agency, UK, for their helpful comments on the bTB model parameters. We also thank Christian Griot from Institute of Virology and Immunoprophylaxis, Mittelhäusern Switzerland, and Heinzpeter Schwermer, Federal Veterinary Office, Switzerland, for their support and constructive discussions related to this project.

REFERENCES

Al-Talafhah, A.H., Lafi, S.Q. and Al-Tarazi, Y. (2003). Epidemiology of ovine brucellosis in Awassi sheep in Northern Jordan. Prev. Vet. Med. 60, 297-306

Hadorn, D.C. and Stärk, K.D.C. (submitted). Evaluation and optimization of surveillance systems for emerging infectious diseases.

Kabagambe, E.K., Elzer, P.H., Geaghan, J.P., Opuda-Asibo, J., Scholl, D.T. and Miller, J.E. (2001). Risk factors for Brucella seropositivity in goat herds in Eastern and Western Uganda. Prev. Vet. Med. 52, 91-108

Kazwala, R.R., Daborn, C.J., Sharp, J.M., Kambarage, D.M., Jiwa, S.F.H. and Mbembati, N.A. (2001). Isolation of Mycobacterium bovis from human cases of cervical adenitis in Tanzania: a cause for concern? Int. J. Tub. Lung Dis. 5 (1), 87-91

Lilienfeld, D.E. and Stolley, P.D. (1994). Foundations of Epidemiology. 3rd edition. New York / Oxford: Oxford University Press

Lithg-Pereira, P.L., Mainar-Jaime, R.C., Alvarez-Sanchez, M.A. and Rojo-Vazquez, F.A. (2001). Evaluation of official eradication-campaigns data for investigating small-ruminant brucellosis in the province of Leon, Spain. Prev. Vet. Med. 51, 215-225

Lithg-Pereira, P.L., Rojo-Vàzquez, F.A. and Mainar-Jaime, R.C. (2004). Case-control study of risk factors for high within-flock small-ruminant brucellosis prevalence in a brucellosis low-prevalence area. Epidemiol. Inf. 132, 201-210

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

Oranga, H.M. and Nordberg, E. (1993). The Delphi panel method for generating health information. Health Policy and Planning 8 (4), 405-412

Slingenbergh, G., Gilbert, M., de Balogh, K. and Wint, W. (2004). Ecological sources of zoonotic diseases. Sci. Tech. Rev. OIE 23 (2), 467-484

Stärk, K.D.C., Regula, G., Hernandez, J., Knopf, L., Fuchs, K., Morris, R.S. and Davies, P. (2006). Concepts for risk-based surveillance in the field of veterinary medicine and veterinary public health: Review of current approaches. BMC Health Services Research. Available from http://www.biomedcentral.com/1472-6963/6/20

SURVEILLANCE FOR CLINICAL PARATUBERCULOSIS: EVALUATION OF ZIEHL-NEELSEN TEST AND ELISA THROUGH BAYESIAN MODELLING

M.F. WEBER*, J. VERHOEFF, G. VAN SCHAIK AND C. VAN MAANEN

SUMMARY

The aim of this study was to evaluate the diagnostic-test characteristics of microscopic examination of Ziehl-Neelsen-stained faecal smears for acid-fast Mycobacteria (ZN-test) and serum-ELISA in cattle suspected of clinical paratuberculosis. Test results of 892 cattle tested simultaneously by ZN-test and serum-ELISA were analysed with Bayesian latent-class models for evaluation of diagnostic tests in two populations without a gold standard. Sampled cattle were divided into two populations in two different ways, based on region and age. Priors for sensitivity and specificity of tests were based on the literature; uninformative priors were used for prevalence's in the various populations. Posterior estimates of sensitivity, specificity, and positive and negative predictive values of the ELISA were always higher than those of the ZN-test, irrespective of the population and choice of model. It is concluded that the ELISA is preferred to the ZN-test to confirm the presumptive diagnosis of clinical paratuberculosis.

INTRODUCTION

Paratuberculosis (Johne's disease) is a chronic inflammatory bowel disease, primarily affecting ruminants. The aetiological agent is *Mycobacterium avium* subsp. *paratuberculosis* (Map). Clinical signs of paratuberculosis in cattle include loss of milk production, weight loss and diarrhoea (National Research Council of the National Academies, 2003)

Testing cattle with clinical signs of paratuberculosis is an important element of surveillance for paratuberculosis. In many infected herds, control of paratuberculosis infection is only initiated after detecting clinical paratuberculosis cases. Diagnostic tests used to confirm a clinical presumptive diagnosis and their characteristics are therefore important, not only in managing the individual patient, but also for paratuberculosis control on the herd-level and on the national level.

The clinical presumptive diagnosis paratuberculosis can be confirmed by demonstrating the presence of Map or presence of antibodies against Map. Methods to demonstrate the presence of Map include faecal culture, microscopic examination of Ziehl-Neelsen-stained faecal smears for the presence of clumps of acid-fast Map organisms (ZN-test) and polymerase chain reaction (PCR) assays (National Research Council of the National Academies, 2003). However, PCR assays are not yet available in the Netherlands for routine testing of large numbers of faecal samples (without prior culturing). Methods to demonstrate antibodies against Map include the

* Maarten F. Weber, GD Animal Health Service, P.O. Box 9, 7400 AA Deventer, The Netherlands. Email: m.weber@gddeventer.com.

complement fixation tests and enzyme-linked immunosorbent assays (ELISA's; Kalis et al., 2002; van Maanen et al., 2002).

To support culling decisions in cases of suspected paratuberculosis, a fast confirmation of the clinical presumptive diagnosis is preferred. Faecal culture, often regarded as a gold standard, takes at least several weeks before a test result is obtained. Therefore, cattle suspected of paratuberculosis are often tested by quicker methods such as ZN-test and serum-ELISA. The sensitivity of the ZN-test has been estimated at 49.3% (Zimmer et al., 1999) in clinically affected cattle. The specificity of the ZN-test was estimated at 83% (Ris et al., 1988). However, only small numbers of cattle were included in these studies. The sensitivity of the serum-ELISA used at the Animal Health Service laboratory, (ELISA Paratuberculosis Antibody Screening, Institut Pourquier, Montpellier, France) has been estimated at 28.0% to 40.8% in faecal culture-positive cattle (van Maanen et al., 2002; Collins et al., 2005). However, to our knowledge, this ELISA has not been evaluated in clinical paratuberculosis cases. Therefore, and because of the rather small sample sizes in the studies on the ZN-test, it is difficult to give a clear advice on the preferred choice of test and interpretation of test-results in cases suspected of clinical paratuberculosis. Therefore, the aim of the present study was to evaluate the diagnostic-test characteristics of microscopic examination of Ziehl-Neelsen stained faecal smears and a serum ELISA in cattle suspected of clinical paratuberculosis.

MATERIALS AND METHODS

Samples

All test results of faecal samples submitted between 1 April 2003 and 1 April 2006 for the ZN-test were retrieved from the laboratory information system of the Animal Health Service. In addition, all ELISA results of serum samples submitted in the same period were retrieved.

Laboratory tests

For the ZN-test, five grams of faeces were suspended in tap water and filtrated through a tea-strainer. The filtrate was mixed with sodiumhypochlorite solution, stored overnight, and centrifuged. A smear was made of the top layer of the sediment, fixated in hot air, stained in a carbolfuchsine solution, decolorized, and then stained in a methylene-blue solution. Samples were considered positive if at least 2 groups of at least 3 acid-fast Mycobacteria were detected at microscopic examination for 10 minutes.

Serum samples were tested with the ELISA Paratuberculosis Antibody Screening [Institut Pourquier, Montpellier, France] according to the manufacturers' instruction. Results of samples with a sample to positive control (S/P) ratio ≤ 0.90 were considered negative, results of samples with an $S/P \geq 1.10$ were considered positive and results of samples with $0.90 < S/P < 1.10$ were considered inconclusive.

Herd-level and animal-level data

For each sample, the unique herd identification number and the animal identification were retrieved, as provided at the submission of the sample. In our analyses, inconclusive test results were assumed to be negative. Each sampled animal was assumed to have been tested only once by ZN-test and at most once by ELISA.

For each herd from which samples were submitted, the region of the country was retrieved in spring 2006. For each sample, attempts were made to uniquely match the identification of the sample provided at submission to a unique cattle identification number in the national cattle identification & registration (I&R) database (Nielen et al., 1996). The date of birth of cattle to which a sample could be matched was retrieved from the I&R database.

Faecal samples submitted for ZN-test and serum samples submitted for ELISA were considered to be collected from the same animal only if (a) collected from cattle in the same herd and (b) the animal identification provided with both samples was exactly the same or both samples could be matched to the same unique cattle identification number, and (c) the difference between the submission dates of both samples was <8 days.

Analysis

To estimate the diagnostic test characteristics (sensitivity, specificity, positive and negative predictive values) of both ZN test and ELISA, Bayesian latent-class models for evaluation of two tests (with or without conditional independence) in two populations described by Branscum et al. (2005) were adapted. Their models are available on-line at http://www.epi.ucdavis.edu/diagnostictests/. Three models were used. In model 1, conditional independence of the ZN-test and ELISA was assumed. In model 2, conditional dependence of the ZN-test and ELISA was assumed. In model 3, conditional dependence of the tests in infected cattle, but conditional independence in non-infected cattle were assumed. The reason for this third model was that, to the authors' knowledge, there is no biologically plausible reason why the tests would be dependent in non-infected cattle (as opposed to infected cattle). The models were run with the freeware program WinBUGS version 1.4.1 (Lunn et al., 2000; available online at http://www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml).

For each of the three models, the population of cattle tested with both tests was divided into two subpopulations with a different prevalence of infection. Known risk factors for clinical paratuberculosis were used to create subpopulations. The incidence of the various differential diagnoses of clinical paratuberculosis was assumed not to be influenced by these risk factors, and to be equally distributed across the subpopulations. Therefore, the proportion of true paratuberculosis cases within the population of cattle showing clinical signs resembling paratuberculosis was expected to be different for the subpopulation exposed to these risk factors and the subpopulation not exposed to these factors. Two risk factors were used to create subpopulations: region of the Netherlands and age at testing. Firstly, the incidence of clinical paratuberculosis is generally thought to be higher in the North of the Netherlands than in other parts of the country, even though there is little published evidence to support this statement. Huitema, (1962) described an annual figure of confirmed clinical diagnoses of 786 in the North, 188 in the East, 240 in the West and 117 in the South of the Netherlands. More recently, the herd-level seroprevalence was found to be higher in the North compared to the East and the West of the Netherlands (Muskens et al., 2000). Therefore, our dataset was divided in a subpopulation of cattle from the North of the Netherlands (i.e. the provinces of Fryslân, Groningen and Drente) and a subpopulation of cattle from the rest of the Netherlands. Secondly, cattle aged ≥ 4 years of age have a 2.7 fold higher incidence of clinical paratuberculosis compared to cattle between 2 and 4 years of age (Reinders, 1963). Similarly, 86 (76%) of 113 cattle with clinical paratuberculosis were ≥ 4 years of age (Benedictus et al., 1987). Therefore, our dataset was divided into subpopulations of cattle <4 year of age and cattle ≥ 4 years of age. Both sets of two subpopulations were analysed separately with the three models described above. Each subpopulation was assumed to include ≥1 clinical case of paratuberculosis.

Uninformative prior distributions (Uniform [0, 1]) were used for the distribution of proportion of paratuberculosis cases within each subpopulation. Because of the low average number of samples per herd (1.2), it was considered appropriate to ignore nesting of samples within herd in our analyses. Furthermore, the size of the population of cattle with clinical signs that may be caused by paratuberculosis is very large, and therefore sampling of cattle was assumed to have been done without replacement.

Prior $\beta(s+1, n-s+1)$ distributions for sensitivity (specificity) of tests were created based on the number s of test-positive (test-negative) and total number n of infected (non-infected) cattle tested in various studies (Table 1).

Table 1. Prior distributions for the sensitivity (Se) and specificity (Sp) of microscopic examination of Ziehl-Neelsen stained faecal smears (ZN) and ELISA.

TEST	PARAMETER	PRIOR DISTRIBUTION	MODE	PERCENTILE OF PRIOR DISTRIBUTION			REFERENCE
				2.5%	50%	97.5%	
ZN	Se	β (38, 39)	0.493	0.383	0.493	0.604	Zimmer et al. (1999)
	Sp	β (16, 4)	0.833	0.604	0.810	0.939	Ris et al. (1988)
ELISA	Se	β (44, 14)	0.768	0.642	0.762	0.859	Egan et al. (1999)
	Sp	β (2133, 4)	0.999	0.996	0.998	0.999	(van Maanen et al., 2002).

Each model was run for $1 \cdot 10^6$ iterations, after a 5000 iterations burn in. Because preliminary analyses indicated some autocorrelation, only every fifth iteration was selected to contribute to the statistics being calculated. For each analysis, history plots, quantile plots and autocorrelation plots were checked for indications of a lack of model fit. With all models, the Gelman-Rubin convergence statistic was monitored in separate runs using two initial strains with initial values for all parameters of 0.001 and 0.999 respectively. An important assumption in the models is that the estimated proportion of true paratuberculosis cases of the two subpopulations is substantially different. This assumption was assumed to be violated if the 95% credibility interval (95% CI) of the posterior difference between the proportions of true paratuberculosis in the two subpopulations included zero. In that case, no further interpretation of the results of the analysis was made.

RESULTS

Descriptive statistics

ZN-test results of 1968 samples were retrieved. In total, 538 (27%) of the 1968 samples tested positive (ZN-positive). For 892 of the 1968 faecal samples from 729 herds, an ELISA result of a serum sample of the same individual was available as well.

In 787 of the 892 cases for which both faecal and serum samples were submitted, one or more clinical signs had been indicated on the submission form: diarrhoea (738 cases), loss of body weight or loss of body condition or wasting (332 cases), loss of milk production (63 cases)

and other clinical signs (30 cases). In 18 cases the presence of clinical signs had been indicated without a specification of these signs. No clinical signs had been indicated on the submission form in the remaining 87 cases. However, because faecal samples are submitted for the ZN-test almost exclusively in case of clinical signs of paratuberculosis, these cases were not excluded from our analyses.

The 892 cattle for which both faecal and serum samples had been submitted originated from 729 herds. Region could be retrieved for all of these herds and cases (Table 2). The cattle identification provided at submission could be matched to a unique animal identification for 665 of the 892 samples. Age at sampling was calculated for each these animals (Table 3).

Table 2. Results of microscopic examination of Ziehl-Neelsen stained faecal smears and ELISA of 892 cattle suspected of clinical paratuberculosis per region of the Netherlands.

REGION	ZN RESULT	ELISA RESULT		TOTAL
		Positive	Negative	
North	Positive	115	8	123
	Negative	130	136	266
Other regions	Positive	135	4	139
	Negative	130	234	364
Total		510	382	892

Table 3. Results of microscopic examination of ZN-stained faecal smears and ELISA of 665 cattle suspected of clinical paratuberculosis per age group.

AGE	ZN RESULT	ELISA RESULT		TOTAL
		Positive	Negative	
< 4 yrs	Positive	42	2	44
	Negative	61	125	186
≥ 4 yrs	Positive	143	6	149
	Negative	134	152	286
Total		380	285	665

Bayesian analysis

In none of the analyses, inspection of history plots, quantile plots, autocorrelation plots and plots of the Gelman-Rubin convergence statistic indicated a lack of model fit. The estimated proportions of true paratuberculosis cases in the various subpopulations are shown in Table 4. In each analysis, the posterior distributions of the proportions of true paratuberculosis cases in the two subpopulations were substantially different, and therefore the results could be interpreted.

Point estimates of the overall proportion of true paratuberculosis cases ranged from 0.604 to 0.674 (Table 4). Models 2 and 3 resulted in slightly higher estimates of the proportion of true paratuberculosis cases than model 1. However, the effect of the choice of parameter used to create subpopulations (region or age) on the estimated overall proportion of true paratuberculosis cases was negligible (Table 4).

The posterior estimates of the overall diagnostic sensitivity, specificity and positive and negative predictive values of the ELISA were always higher than those of the ZN-test. In analyses with subpopulations based on age, point estimates of the overall negative predictive value of ZN-test and ELISA ranged from 0.457 to 0.550 and 0.740 to 0.891, respectively (Table 5). Point estimates of the overall positive predictive value of the ZN-test ranged from 0.974 to 0.976. In contrast, the point estimates of the overall positive predictive values of the ELISA were always 0.999 (Table 5). Similar results were obtained with subpopulations based on region (results not shown).

The estimates of the specificities and positive predictive values of the tests were almost identical across models 1, 2 and 3 with subpopulations based on age (Table 5). This was related to the identical probabilities with model 2 of a positive ELISA result in ZN-positive and ZN-negative non-diseased cattle (Table 6), i.e. the tests were conditionally independent in non-diseased cattle. However, estimates of the sensitivities and negative predictive values of the tests were somewhat lower with models 2 and 3 (conditional dependence of tests in diseased cattle) compared to model 1 (conditional independence of tests; Table 5). Similar results were obtained with subpopulations based on region (results not shown).

Table 4. Estimated proportions (95% CI) of cattle with paratuberculosis within subpopulations of 892 cattle with clinical signs resembling paratuberculosis. For each subpopulation, the sample size n is indicated between brackets.

SUBPOPULATIONS BASED ON	MODEL	PROPORTION (95% CI) OF CATTLE WITH PARATUBERCULOSIS			
		OVERALL	WITHIN SUBPOPULATION		DIFFERENCE BETWEEN SUBPOPULATIONS
Region		(n= 892)	North (n=389)	Other regions (n=503)	
	Model 1	0.604 (0.569, 0.640)	0.671 (0.619, 0.722)	0.553 (0.506, 0.600)	0.118 (0.050, 0.185)
	Model 2	0.674 (0.616, 0.754)	0.748 (0.674, 0.844)	0.618 (0.552, 0.700)	0.131 (0.055, 0.208)
	Model 3	0.651 (0.603, 0.710)	0.723 (0.658, 0.796)	0.596 (0.539, 0.661)	0.127 (0.053, 0.200)
Age		(n=665)	< 4 yrs (n=230)	≥ 4 yrs (n=435)	
	Model 1	0.605 (0.565, 0.647)	0.473 (0.406, 0.542)	0.676 (0.626, 0.725)	-0.203 (-0.285, -0.120)
	Model 2	0.673 (0.611, 0.757)	0.527 (0.445, 0.620)	0.751 (0.679, 0.847)	-0.226 (-0.321, -0.133)
	Model 3	0.650 (0.597, 0.713)	0.508 (0.432, 0.590)	0.725 (0.663, 0.798)	-0.218 (-0.308, -0.129)

Table 5. Estimated sensitivity (Se), specificity (Sp), negative predictive value (PVN) and positive predictive value (PVP) of microscopic examination of ZN-stained faecal smears (ZN) and ELISA in analyses using subpopulations by age (<4 yrs versus ≥4 yrs of age).

TEST	TEST CHARACTERISTIC	MODEL		
		MODEL 1	MODEL 2	MODEL 3
ZN	Se	0.477 (0.431, 0.524)	0.436 (0.380, 0.489)	0.476 (0.430, 0.521)
	Sp	0.980 (0.956, 0.994)	0.976 (0.946, 0.993)	0.978 (0.951, 0.993)
	PVN	0.550 (0.498, 0.598)	0.457 (0.341, 0.538)	0.502 (0.416, 0.568)
	PVP	0.974 (0.941, 0.992)	0.975 (0.944, 0.992)	0.976 (0.947, 0.993)
ELISA	Se	0.920 (0.880, 0.951)	0.829 (0.750, 0.888)	0.856 (0.796, 0.903)
	Sp	0.998 (0.996, 1.000)	0.998 (0.996, 1.000)	0.998 (0.996, 1.000)
	PVN	0.891 (0.832, 0.935)	0.740 (0.570, 0.843)	0.789 (0.673, 0.866)
	PVP	0.999 (0.997, 1.000)	0.999 (0.998, 1.000)	0.999 (0.997, 1.000)

Table 6. Median (95% CI) posterior estimates of probabilities of a positive and negative ELISA result given disease status and result of the ZN-test (i.e. model parameters related to conditional dependence of tests) in models 2 and 3 with subpopulations based on age.

PARAMETER	MODEL	
	MODEL 2	MODEL 3
P(ELISA-positive | diseased, ZN-negative)	0.759 (0.641, 0.856)	0.799 (0.700, 0.879)
P(ELISA-positive | diseased, ZN-positive)	0.921 (0.881, 0.951)	0.920 (0.881, 0.951)
P(ELISA-negative |non-diseased, ZN-positive)	0.998 (0.996, 1.000)	N.A.
P(ELISA-negative |non-diseased, ZN-negative)	0.998 (0.996, 1.000)	N.A.

N.A.: not applicable.

DISCUSSION

The results of this study show that the posterior estimates of the sensitivity, specificity and positive and negative predictive values of the ELISA were significantly higher than those of the ZN-test. Therefore the ELISA appears to be superior to the ZN-test to confirm the presumptive diagnosis of clinical paratuberculosis. The positive predictive value of the ELISA was estimated at 0.999 (0.997, 1.000). This means that very little diagnostic information can be gained with the ZN-test if the ELISA has a positive result. Also, if the ELISA has a negative result, the likelihood of gaining diagnostic information with the ZN-test is very small, because this test was positive in only 3% of ELISA-negative cases (Table 2). Furthermore, the ZN-test is very laborious, can not be automated and is expensive in comparison to the ELISA. The results of this study indicate that the ELISA alone is a cost-effective test to confirm the presumptive diagnosis of clinical paratuberculosis.

The posterior proportion of true paratuberculosis cases within the population of cattle suspected of paratuberculosis was higher in the North of the Netherlands compared to other regions. This is in line with our hypothesis that the incidence of clinical paratuberculosis is higher in the North, whereas other causes of clinical signs resembling paratuberculosis may be equally distributed across regions. Similarly, the posterior proportion of true paratuberculosis cases was higher in older cattle than in younger cattle suspected of paratuberculosis.

The posterior estimates of the sensitivity of the ZN-test were comparable to previously published estimates of 49.3% in clinically affected cattle (Zimmer et al., 1999) and 56% in faecal culture-positive cattle (Ris et al., 1988). This is, of course, partially explained by our prior for the sensitivity that was based on these studies. However, the posterior estimates of specificity (>95%) in this study were higher than the estimate of 0.83 (95% CI: 0.59, 0.96) of Ris et al. (1988). This may be related to the strict criterion used to declare a sample positive in the ZN-test in the present study: samples were only considered positive if at least 2 groups of at least 3 acid-fast Mycobacteria were detected at microscopic examination of smears. Other factors that may have contributed to the high specificity were a conscientious examination of smears by experienced laboratory technicians in this study and, possibly, a lower background population of Mycobacteria compared to the study by Ris et al. (1988). The posterior estimates of the sensitivity of the ELISA in this study were ≥80%, broadly in line with estimates in clinical cases of 77% by Egan et al. (1999) and 87% by Sweeney et al. (1995), but higher than the estimate of 50% by Bech-Nielsen et al. (1992). The posterior estimate of the specificity of the ELISA used in this study was in line with previously published estimates (van Maanen et al., 2002; Collins et al., 2005) which is related to the fact that a highly informative prior was used, based on the study of van Maanen et al. (2002).

In our study, three models were used, with and without the assumption of conditional independence of the tests. Conditional dependence of tests resulted in somewhat lower estimates of the sensitivities and positive predictive values. However, the effect of this assumption on the posterior distributions of the specificity and negative predictive values was negligible. Moreover, the assumption of conditional dependence of tests had no practical consequences for the preferred choice of test, because with each of the models, the test-characteristics of the ELISA were more attractive than those of the ZN-test. Prior to the analyses, conditional dependence of the ZN-test and ELISA was hypothesized to be plausible in infected cattle, but not in uninfected cattle. Therefore, model 3 was considered the biologically most plausible

model. This hypothesis is supported by the results of model 2, indicating conditional dependence of tests in infected cattle but independence of tests in uninfected cattle.

Constant test accuracy across the subpopulations within an analysis was assumed in the present study. At first sight, this assumption may seem in conflict with observed associations between diagnostic sensitivity of tests for paratuberculosis and age (Jubb et al., 2004; Nielsen and Toft, 2006). However, these observed associations are likely to be the result of an association between age and the stage of the infection-and-disease process on the one hand and an association between stage of the infection-and-disease process and sensitivity on the other hand. In the present study, only the final stage of the infection-and-disease process, i.e. clinical disease, was studied. Therefore, the assumption of constant test accuracy across subpopulations, including subpopulations based on age, was considered not to be violated. However, even if test accuracy would differ between subpopulations, our analysis still applies. Then the resulting estimates can be interpreted as average values across both populations rather than population-specific values, which would still be of interest (Branscum et al., 2005).

It is concluded that the sensitivity, specificity, and positive and negative predictive value of the ELISA in cattle suspected of clinical paratuberculosis are higher than those of the ZN-test. Therefore, to confirm the presumptive diagnosis of paratuberculosis, the ELISA is preferred above the ZN-test. If the ELISA is used to confirm this presumptive diagnosis, little diagnostic information can be gained by performing the ZN-test as well.

ACKNOWLEDGEMENTS

The authors thank I. Gardner for his advice on the Bayesian models prior to the study, F. de Vries for providing the data on soil types, A. Luppen and H. Assink for assistance in data collection, and W. Swart for statistical advice.

REFERENCES

Bech-Nielsen, S., Jorgensen, J.B., Ahrens, P., Feld, N.C., 1992. Diagnostic accuracy of a *Mycobacterium phlei*-absorbed serum enzyme-linked immunosorbent assay for diagnosis of bovine paratuberculosis in dairy cows. J. Clin. Microbiol. 30, 613-618.

Benedictus, G., Dijkhuizen, A.A., Stelwagen, J., 1987. Economic losses due to paratuberculosis in dairy cattle. Vet. Rec. 121, 142-146.

Branscum, A.J., Gardner, I.A., Johnson, W.O., 2005. Estimation of diagnostic-test sensitivity and specificity through Bayesian modeling. Prev. Vet. Med. 68, 145-163.

Collins, M.T., Wells, S.J., Petrini, K.R., Collins, J.E., Schultz, R.D., Whitlock, R.H., 2005. Evaluation of five antibody detection tests for diagnosis of bovine paratuberculosis. Clin. Diagn. Lab Immunol. 12, 685-692.

Egan, J., Weavers, E., O'Grady, D., 1999. An evaluation of diagnostic tests for Johne's disease in cattle. Irish Veterinary Journal 52, 86-89.

Huitema, H., 1962. Diagnose en prognose bij paratuberculose van het rund. PhD thesis, Utrecht University, 227 pp.

Jubb, T.F., Sergeant, E.S., Callinan, A.P., Galvin, J., 2004. Estimate of the sensitivity of an ELISA used to detect Johne's disease in Victorian dairy cattle herds. Aust. Vet. J. 82, 569-573.

Kalis, C.H.J., Barkema, H.W., Hesselink, J.W., van Maanen, C., Collins, M.T., 2002. Evaluation of two absorbed enzyme-linked immunosorbent assays and a complement fixation test as replacements for fecal culture in the detection of cows shedding *Mycobacterium avium* subspecies *paratuberculosis*. J. Vet. Diagn. Invest 14, 219-224.

Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS -- a Bayesian modelling framework: concepts, structure, and extensibility. Statistics and Computing, 10, 325-337.

Muskens, J., Barkema, H.W., Russchen, E., Van Maanen, K., Schukken, Y.H., Bakker, D., 2000. Prevalence and regional distribution of paratuberculosis in dairy herds in The Netherlands. Vet. Microbiol. 77, 253-261.

National Research Council of the National Academies, 2003. Diagnosis and Control of Johne's disease. National Academies of Sciences, Washington, D.C., 229 pp.

Nielen, M., Jansen, F.C., van Wuijckhuise, L.A., Dijkhuizen, A.A., 1996. [I&R (identification and registration) system cattle: an analysis of its use during a foot-and-mouth-disease outbreak in The Netherlands]. Tijdschr. Diergeneeskd. 121, 576-581.

Nielsen, S.S., Toft, N., 2006. Age-specific characteristics of ELISA and fecal culture for purpose-specific testing for paratuberculosis. J. Dairy Sci. 89, 569-579.

Reinders, J. S., 1963. Bestrijding van klinische paratuberculose bij runderen. PhD thesis, Utrecht University, 194 pp.

Ris, D.R., Hamel, K.L., Ayling, J.M., 1988. The detection of *Mycobacterium paratuberculosis* in bovine faeces by isolation and the comparison of isolation with the examination of stained smears by light microscopy. N. Z. Vet. J. 36, 112-114.

Sweeney, R.W., Whitlock, R.H., Buckley, C.L., Spencer, P.A., 1995. Evaluation of a commercial enzyme-linked immunosorbent assay for the diagnosis of paratuberculosis in dairy cattle. J. Vet. Diagn. Invest 7, 488-493.

van Maanen, C., Koster, C., van Veen, B., Kalis, C.H.J., Collins, M.T., 2002. Validation of *Mycobacterium avium* subsp. *paratuberculosis* antibody detecting ELISA's. In: Juste, R.A., Geijo, M.V., Garrido, J.M. (Eds.), Proceedings of the 7th International Colloquium on Paratuberculosis, Bilbao, Spain, 11-14 June 2002, p. 182.

Zimmer, K., Drager, K.G., Klawonn, W., Hess, R.G., 1999. Contribution to the diagnosis of Johne's disease in cattle. Comparative studies on the validity of Ziehl-Neelsen staining, faecal culture and a commercially available DNA-Probe test in detecting Mycobacterium paratuberculosis in faeces from cattle. Zentralbl. Veterinarmed. B 46, 137-140.

RISK-BASED SAMPLING VERSUS RANDOM SAMPLING FOR MONITORING TETRACYCLINE RESIDUES IN SLAUGHTERED CALVES IN SWITZERLAND

P. PRESI, K.D.C. STÄRK, L. KNOPF, E. BREIDENBACH, M. SANAA, J. FREY AND G. REGULA*

SUMMARY

In European countries, chemical residues are routinely monitored in slaughtered animals. EU requirements advise member countries partially to target their sampling in order to increase the likelihood of finding residues. The objective of this study was to assess the efficacy of a risk-based sampling approach compared with random sampling for the monitoring of tetracycline residues in calves. A stochastic model including real data as well as expert opinion was developed. It demonstrated that the risk-based approach increased the efficacy of the monitoring programme by up to 100% compared with random sampling. The greatest benefit of risk-based sampling was observed when the prevalence of contamination was low in the population. The resources saved through the risk-based sampling can be used to improve the quality of the information on risk factors. With this information, the model can be further improved as better estimates of the parameters become available.

INTRODUCTION

Switzerland has been conducting a monitoring programme for selected contaminants in animal-derived food since 1996. The aim is to ensure a safety level for the consumer of meat or meat products according to Swiss and European Union legislation (Council Directive 96/23/EC). While human and financial resources available to government veterinary services are becoming scarcer in many countries, a high level of safety is expected for the consumer of meat and meat products. Therefore priorities need to be set.

EU requirements advise member countries partially to target their sampling in order to increase the likelihood of finding residues (EC 802/2004). This approach defined as risk-based monitoring (Stärk et al. 2006) is used in many different fields, e.g. veterinary epidemiology (Hadorn et al. 2002), human health (Payne-Sturges et al. 2004) and ecology (EPA 1995). The main goal of a risk-based approach is to achieve a higher benefit-cost ratio with existing or reduced resources, i.e. ensure an equivalent level of safety for a lesser investment of resources. Although the current Swiss residue monitoring programme is stratified and partially risk-based, a detailed assessment of the sampling showed that potential risk factors were not adequately represented in the sample (Presi et al. 2007a).

* Gertraud Regula. Federal Veterinary Office. Schwarzenburgstr. 155, CH – 3097 Liebefeld, Switzerland. Email: Gertraud.Schuepbach@bvet.admin.ch

The main objective of this study was to assess the efficacy of a risk-based sampling approach compared with random sampling for monitoring tetracycline residues in calves. This combination of substance and animal type was used for the evaluation of risk-based sampling, because of the quality of data available and the widespread use of this substance in Switzerland (Anonymous 2006b). To achieve this assessment, a stochastic model was developed to simulate the two sampling strategies, risk-based and random (Presi et al. 2007b).

The specific objectives were: (1) to define the minimum sample size needed in both risk-based and random sampling, in order to detect at least one contaminated herd assuming a defined prevalence of contaminated herds in the baseline population; (2) to define the minimum prevalence of contaminated herds which can be detected with the sample size specified in EU requirements (approximately 120 samples), and show the number of positives that can be detected in both random and risk-based sampling; (3) to assess the influence of the different input parameters on the output of the model.

MATERIALS AND METHODS

The model developed for the evaluation of the efficacy of risk-based sampling compared with random sampling was structured as follows. Starting with a baseline population and expert opinion, a certain percentage of the calf herds was simulated to be contaminated with tetracycline residues. Using real data from a past monitoring programme, risk factors on the presence of residues were extracted in order to perform a risk-based sampling on all herds, including simulated-positive ones. This sample was then compared with random sampling of the same population. Further details on the model structure are available (Presi et al. 2007b).

Data description

From the 47 842 farms present in Switzerland in 2005, 26 967 farms producing calves formed the baseline population for this model. The database available from the Federal Office for Agriculture contained the following factors that could potentially be associated with a farm producing tetracycline contaminated calves:

- *bio* Farms that follow regulations on organic farming in comparison with conventional farms. This variable is binary (0 = conventional, 1 = organic). Twelve percent of the farms were organic.
- *bts* Farms that receive subsidies according to the edict of the Federal department of economic affairs for animal-friendly husbandry systems (RS 910.132.4). This variable is binary (0 = don't get subsidies, 1 = get subsidies). Forty-two percent of the farms got bts subsidies.
- *raus* Farms that receive subsidies according to the edict of the Federal department of economic affairs for letting the animals outside on a regular basis (RC 910.132.5). This variable is binary (0 = don't get subsidies, 1 = get subsidies). Seventy-five percent of the farms got raus subsidies.
- *herdsize* Number of animals present on the farm for each livestock category (bovine - calves), on a yearly basis. The bovine category includes the calves. These variables are continuous. Bovine herdsize (calves included) ranged from one to 804 animals, and calves herdsize ranged from one to 395 animals.
- *nb of species* Number of species present on the farm. This variable is binary (0 = more than one species; 1 = one species). Seventy-eight percent of the farms carried only one species.

- *LAU* (large animal unit) Number of animals present on the farm and weighted according to the livestock categories. This variable is continuous. The farm sizes ranged from 0.339 to 610 large animal units.
- *mvmt* Number of movements observed to a farm resulting in arrival of new animals in the herd per animal, on a yearly basis. This variable is continuous. The number of bovine movements to a farm per year ranged from 0.015 to 9.27 movements per bovine.
- *field-size per LAU* Number of hectares of field that are available per LAU on the farm. This variable is continuous. The field-size ranged from 0.03 to 1.44 hectares per large animal unit.

Simulation of contaminated herds

Five experts were asked to quantify the importance of a list of potential risk factor for explaining the presence of residues of tetracycline in calves. The risk factors were identified from the literature. The experts stated direction and size of the association (e.g. how strong is the relationship between the farm type (organic vs. conventional) and the probability that a herd is treated with tetracycline). For each variable a relative risk score was defined from the consensus value given by the five experts. In the case of a binary variable both parameters would sum to one (e.g. 0.65 for conventional farms compared to 0.35 for organic farms). For continuous variables, the parameters would increase or decrease (depending on the direction of the relationship) proportionally to the increase of the factors. The minimum and the maximum value would sum to one (e.g. 0.1 for the lowest number of calves on the farm and 0.9 for the highest number of calves on the farm). Values above the 95% percentiles were attributed to the upper or lower value of the factor to decrease the effect of outliers. The scores were multiplied to determine the relative risk score of each farm sending calves with tetracycline residues to slaughter. In the model, this score was used as a probability to define which herds would be contaminated with tetracycline residues.

Design of risk-based sampling

In one large Swiss laboratory, an intensive monitoring programme of residues in urine is conducted every year involving animals slaughtered in two major slaughterhouses in Switzerland. Urine samples collected from 602 calves at slaughter were analysed for the presence of tetracycline residues in the years 2004 to 2005. Positive samples were defined as samples where residues of tetracycline could be detected. These samples were traced back to their herd of origin. A herd with at least one positive animal was defined as being positive. Samples originated from 398 different herds, and 136 (34%) of these herds were positive for residues of tetracycline. The risk-based sampling protocol was developed by analysing the possible risk factors (as described in section 'data description') for the 398 farms with results on tetracycline residues in urine. The regression parameters and their variance were estimated using generalized estimation equations. For each herd in Switzerland, a probability of sampling was calculated using the equation of the final generalized linear regression model. Thus, herds for which the regression model predicted a high risk of tetracycline residues had a high probability of being included in the risk-based sample.

Stochastic model

The stochastic model integrating the data on all Swiss calf herds, expert opinion, and the regression equation for risk-based sampling was performed using the language and environment for statistical computing "R" (R development Core Team, 2006). A total of 10 000 iterations were run for each of seven different pre-set prevalences of contaminated herds: 0.5%, 1%, 2%,

5%, 10%, 20% and 50%. One sampling method consisted of sampling herds until the first contaminated samples were detected. The second objective was to sample 120 herds and to count how many contaminated samples were found.

A sensitivity analysis was performed to explore how the outputs of the model were influenced by the quality of the input parameters. A correlation matrix was defined to quantify the importance of each potential risk factor on the probability of a herd being contaminated. To evaluate the influence of the data from the past urine monitoring programme, a bootstrap approach was used. A total of 100 iterations were performed re-sampling each time 398 samples with replacement out of the 398 results of analyses from the past urine monitoring programmes. For each iteration, the risk factors were re-calculated from the generalized linear model. Then, 100 iterations were run in the model with a prevalence of 2% for each combination of risk factors. This allowed an estimate to be made of the variability induced by the risk factors.

RESULTS

Expert opinion

According to expert opinion, the following variables were classified as being associated with the presence of tetracycline residues in Swiss calf herds (Table 1). The experts ranked the herd size (number of calves) and the frequency of animal movements as most important for the risk of tetracycline residues. By multiplying the relative-risk scores of all variables for each farm in Switzerland, an individual value was obtained per herd showing the relative risk score of this particular farm having calves contaminated with residues compared with the other farms. The higher limit of the 95% CI for the risk score differed by a factor 60 from the lower limit. These results were used to simulate the level of contamination in the baseline farm population as described above.

Risk factor analysis

Significant parameters obtained from the logistic regression associated with the presence of residues of tetracycline in urine samples were the number of calves present in the herd and the number of large animal units on the farm (Table 2). The Pearson correlation coefficient between the two parameters was 0.13. At the cut-off for the prediction of the risk of tetracycline residues the regression model predicted 136 herds to be positive for tetracycline residues, 72 were actually contaminated (specificity 53%). Out of 262 herds predicted to be negative for the presence of residues of tetracycline, 200 were actually negative (sensitivity 76%). For each farm, a relative-risk of tetracycline residues was derived from the generalized linear model. These results were used to target the risk-based sampling as described above.

Table 1. Potential risk factors associated with the presence of residues in calf herds according to the experts' opinion in Switzerland. A larger difference between the minimum and maximum relative risk scores indicates a greater importance of the parameter associated with the presence of tetracycline residues.

NAMES	FACTOR VALUE		RELATIVE SCORE	
	MIN	MAX	MIN	MAX
Number of bovine movements per year	0.015	≥0.88	0.10	0.90
Number of calves on the farm	1	≥24	0.10	0.90
Number of species on the farm	1	>1	0.30	0.70
Number of bovines on the farm	1	≥87	0.30	0.70
Large animal units (*LAU*)	0.35	≥60	0.30	0.70
Field-size per *LAU*	0.03	≥1.41	0.30	0.70
Farm-type	organic	conventional	0.35	0.65

Table 2. Significant parameters in the logistic regression analysis of tetracycline residues in 398 urine samples of calves sampled in Switzerland

FACTORS	ESTIMATE	95% CI	P-VALUE
Number of calves on the farm	0.046	[0.036;0.056]	<0.001
Number of large animal units	0.021	[0.014;0.028]	0.003

Comparison of random vs. risk-based sampling

The sample size for detecting at least one contaminated herd in the sample was determined by running 10 000 iterations for prevalence scenarios of 0.5%, 1%, 2% and 10% respectively. At a prevalence of 0.5% of herds with tetracycline residues, a median number of 97 farms (95% CI [4;510]) needed to be sampled in the risk-based sampling programme compared to 135 farms (95% CI [6;713]) for random sampling. At a prevalence of 1% the risk-based sampling resulted in a median value of 47 farms (95% CI [2;255]) versus 69 (95% CI [3;358]) for random sampling. At a prevalence of 2% the median values were 25 (95% CI [1;131]) and 35 (95% CI [2;180]) for risk-based and random sampling respectively. A prevalence of 10% resulted in median values of 5 (95% CI [1;26]) and 7 (95% CI [1;35]) for the two sampling methods.

When the sample size was fixed at 120 herds, risk-based sampling detected a median of one contaminated herd (95% CI [0;3]) versus zero for random sampling (95% CI [0;2]) at a prevalence of 0.5%. The median number of detected positive herds with the risk-based and random sampling protocols were two (95% CI [0;5]) versus one (95% CI [0;4]), three (95% CI [0;7]) versus two (95% CI [0;6]) and 16 (95% CI [9;24]) versus 12 (95% CI [6;19]) for prevalences of 1%, 2%, and 10% respectively.

At a prevalence of 0.5%, the risk-based sampling detected at least one contaminated herd in 58.8% of the simulations with a fixed sample size of 120. This percentage was 45% for random sampling. At a prevalence of 2%, the risk-based sampling detected at least one contaminated

herd in 97.1% of the simulations versus 91.1% for the random sampling. At a prevalence of 5%, risk-based detection reached 99.9% versus 99.8% for random sampling.

Varying the prevalence of contaminated herds had a direct influence on the performance of the risk-based sampling compared with random sampling (Fig. 1). At a prevalence of 1% of contaminated herds, the risk-based sampling increased the efficacy of the sampling by 100% compared with random sampling. At a prevalence of 2% of contaminated herds, 50% more positive herds were detected with risk-based compared with random sampling, whereas the increase in efficacy was only 12% at a prevalence of 50%.

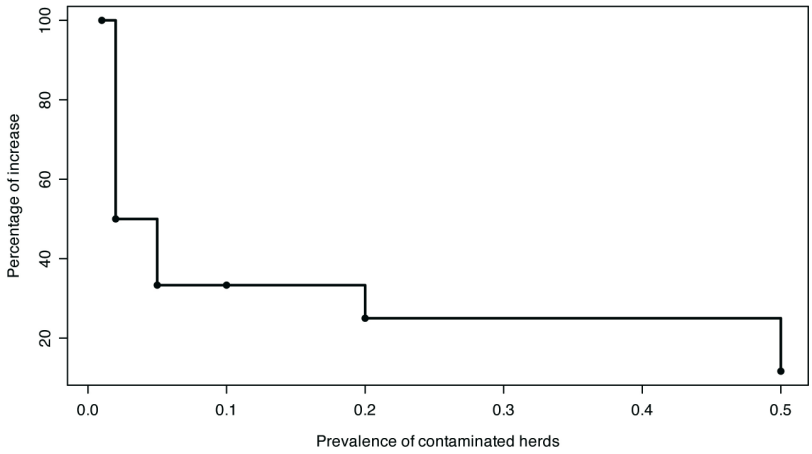


Fig. 1 Increase in percentage of positive samples detected using risk-based sampling when compared with random sampling in relation to the prevalence of herds contaminated with tetracycline in a simulated baseline population.

The correlation coefficient between the potential risk factors defined by the experts and the probability of contamination of a herd was 0.72 for the number of calves present on the farm, 0.58 for the total number of bovines, 0.42 for the number of movements, 0.35 for the number of large animal units, 0.16 for the number of species present on the farm, 0.14 for the farm type and 0.07 for the field-size per LAU on the farm.

The bootstrap approach gave results consistent with those obtained by running the model with fixed risk factors from the generalized linear model (data not shown). Using regression equations from bootstrapping to derive the risk-based sampling resulted in identical median values for the number of herds which had to be sampled to detect one positive, and in the number of positive herds detected in a sample of 120.

DISCUSSION

The model demonstrated that the risk-based approach could increase the efficacy of the monitoring programme by up to 100% compared with simple random sampling. This gain in efficacy should allow for further improvement of residue monitoring. If the sample size used in the current random sampling is maintained, this should lead to the detection of larger numbers of contaminated samples, and thus to improved consumer protection. Through further detailed analysis of the herd characteristics of these samples, the risk factors could be refined and complemented with other more informative ones, leading to a further increase in efficacy of the

risk-based sampling. This is illustrated in the study of Kuntz et al. (2003), who could increase the efficacy of their targeted sampling protocol, using extra sources of information obtained from local knowledge. The difference in efficacy between risk-based and random sampling was directly correlated with the prevalence of contaminated herds. The greatest benefit of risk-based sampling was seen when the residue prevalence in the population was low. This finding is consistent with the results of a study on the benefit of risk-based sampling for detection of paratuberculosis infection in dairy herds (Tavornpanich et al. 2006). As the effort for obtaining accurate prior information on risk factors is similar for a high or low prevalence of residues, it is much more rewarding to perform a risk-based monitoring with a low prevalence.

To evaluate the efficacy and the repeatability of the two sampling strategies used in the model, a probability of occurrence of tetracycline residues had to be defined for each Swiss calf herd. Expert opinion was used to model the contaminated herds. The distribution of the relative risk scores for herds being positive for tetracycline residues was log normal. Therefore a few herds had the highest probability of being contaminated with residues. This is consistent with the results obtained in the frame of the monitoring programme in Switzerland where zero positive samples were found in 2004, two in 2005 and 16 in 2006 (Anonymous 2006a). The majority of farms use the correct withdrawal period after tetracycline application before slaughtering. The expert opinion procedure allowed an estimate for the risk factors even though few data were available. By using five experts from different backgrounds and in an independent setting, we attempted to minimise bias in the interpretation of the parameters. In addition, past data from a monitoring programme were integrated in the model to define the risk factors for the risk-based sampling. All significant risk factors in the logistic regression were also evaluated as important by the experts, even though none of them knew the data from the past monitoring programme. This illustrated that expert opinion was plausible for the Swiss calf production. Nevertheless it should be stated that this situation might differ in other countries.

The potential effect of expert opinion on the output of the model was assessed using the correlation coefficient between the value of the risk factors as defined by the experts and the probability of contamination derived from those factors. The number of calves in the herd as well as the number of bovines had a correlation of 0.72 and 0.58 respectively, with the predicted presence or absence of residues. The effect of possible bias in the expert opinion would be greatest for these factors. Nevertheless, experts' estimates of the effect of herd size were confirmed by the results of the logistic regression analysis. The outputs of the model (sample-size) allowed for evaluating the repeatability of both sampling schemes (Thurmond 2003). The consistency of the outputs is a measure of the precision of a sampling scheme. Looking at both ranges of the confidence intervals obtained between the risk-based and the random approach, the risk-based approach tended to have a smaller variance, showing a greater precision than the random approach. By improving the quality of prior information, the consistency could be further increased.

The poor predictive value, as well as the low specificity of the generalized linear regression model defined to explain the presence of tetracycline residues in urine samples, showed that the risk factors available for the analysis are probably not be the most important factors determining the risk of tetracycline residues. For instance, more informative data would be the amount of tetracycline prescribed in a herd. This study illustrated the need for information of higher quality to increase the predictive power of the model, and thus optimise the risk-based sampling. However, it could be demonstrated that sampling error of the residue monitoring programme did not affect the output of the model, as the bootstrap procedure gave consistent results with fixed parameters for the generalized linear model.

Urine samples were used for the analyses because the level of detection is much lower than in meat, even though for consumer protection, residues in meat are of main interest. This could potentially induce bias if the risk factors for residues in meat differed from the risk factors for residues in urine. Residues in urine and meat are both caused by tetracycline application to the animals, even though urine may also test positive for residues if the withdrawal period is applied correctly (Korsrud et al. 1996).

The model demonstrated that efficacy was improved by switching from a random sampling to a risk-based sampling. The performance of risk-based sampling depends greatly on the quality of data on risk factors. To start a risk-based sampling as part of a monitoring programme, an investment of resources with the aim of obtaining prior information needs to be considered. Even with a limited amount of information in combination with expert opinion, a risk-based sampling approach can be designed that will improve the results obtained with random sampling. The resources saved through the risk-based sampling can be used to improve the quality of prior information. With this information, the model can be further improved as better estimates of the parameters become available.

ACKNOWLEDGEMENTS

The authors would like to thank Corinne Baehler, Thomas Kaufmann, Christian Casura, Peter Jakob and Clemens Bauer for the quality of their inputs. The authors also thank Bruno Paciarelli and the cantonal Labor of Zurich for the quality of their support and the quality of the data provided. Finally the authors also thank Sébastien Nusslé for his inputs regarding the modelling.

REFERENCES

Anonymous (2006a). Annual report 2006. In Swiss federal veterinary office magazine, office, S.f.v., ed. (Bern, Swiss federal veterinary office), 26-33

Anonymous (2006b). Distribution d'antibiotiques utilisés en médecine vétérinaire : comparaison des données 2004 et 2005 (Bern, Swissmedic, Institut suisse des produits thérapeutiques), 4

Council Directive 96/23/EC. Available: http://europa.eu.int/eurlex/lex/LexUriServ/LexUriServ.do?uri=CELEX:31996L0023:EN:HTML

EPA (1995). Use of risk-based decision-making in UST Corrective Action Programs (Washington D.C., U.S. Environmental Protection Agency), 15

Hadorn, D.C., Rufenacht, J., Hauser, R. and Stärk K.D. (2002). Risk-based design of repeated surveys for the documentation of freedom from non-highly contagious diseases. Prev. Vet. Med. 56, 179-192

Korsrud, G.O., Papich, M.G., Fesser, A.C., Salisbury, C.D. and MacNeil, J.D. (1996). Residue depletion in tissues and fluids from swine fed sulfamethazine, chlortetracycline and penicillin G in combination. Food Add. and Cont. 13, 287-292

Kuntz, R.L., Hartel, P.G., Godfrey, D.G., McDonald, J.L., Gates, K.W. and Segars, W.I. (2003). Targeted sampling protocol as prelude to bacterial source tracking with Enterococcus faecalis. J. Environ. Qual. 32, 2311-2318

Payne-Sturges, D.C., Burke, T.A., Breysse, P., Diener-West, M. and Buckley, T.J. (2004). Personal exposure meets risk assessment: a comparison of measured and modeled exposures and risks in an urban community. Environ. Health 112, 589-598

Presi, P., Knopf, L., Regula, G., Pacciarelli, B., Frey, J. and Stärk, K.D. (2007a). Evaluation of the chemical residue monitoring in animal-derived products in Switzerland. Food Add. and Cont. 24, 590-597

Presi, P., Stärk, K.D., Knopf, L., Breidenbach, E., Sanaa, M., Frey, J. and Regula, G (2007b). Efficiency of risk-based sampling versus random sampling for monitoring tetracycline residues in slaughtered calves in Switzerland. Food Add. and Cont. *In press*

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

Tavornpanich, S., Gardner, I.A., Carpenter, T.E., Johnson, W.O. and Anderson, R.J. (2006). Evaluation of cost-effectiveness of targeted sampling methods for detection of Mycobacterium avium subsp paratuberculosis infection in dairy herds. Am. J. Vet. Res. 67, 821-828

Thurmond, M.C. (2003). Conceptual foundations for infectious disease surveillance. J. Vet. Diag. Invest. 15, 501-514

FOOT AND MOUTH DISEASE

RISK-CLASSIFICATION OF FINNISH PIG FARMS BY SIMULATED FMD SPREAD

T. LYYTIKÄINEN* AND E.R. KALLIO

SUMMARY

Farms may be classified based on their capability to spread infectious diseases. This could be applicable for instance in targeting disease surveillance to the farms which are most efficient in spreading disease. The spread of an infectious disease in the Finnish pig farm population was simulated using a Monte Carlo simulation model. Individual farms were used as the starting point of foot and mouth disease epidemics for a large number of iterations. The farm specific information on pig transportation and farm location were used to give a partial description of the contact network of the farm. The risk-classification of the farms was then derived from the simulation outputs. The highest risk-class contained a small proportion (10%) of the farms. The results indicated that farms in the highest risk-class tended to become infected from other farms in the highest risk-class and therefore they were strongly interconnected. A developed ordinal regression model predicted risk-classes quite well.

INTRODUCTION

Classification of pig farms, based on the risk of spreading infectious diseases to other farms or becoming infected by other farms, may be utilized for risk management purposes (Mintiens et al., 2003). Farm classification, based on the possible influence they may have on disease epidemics, may help to allocate finite resources efficiently in disease surveillance and monitoring (Stärk et al., 2006). To risk-classify the farms, it is essential to recognise measurable farm level characteristics, i.e. the factors which will predict the risk they pose to other farms (Mintiens et al., 2003; Stärk et al., 2006). The risk-classification of pig farms can be based on several factors such as production type or herd size.

Often farms are classified based on type of production. Firstly, farms may be divided on the basis of the species they breed. Secondly, pig farms, for instance, can be divided into sow farms, finishing farms and mixed farms. In addition, there may be artificial insemination stations, performance test stations, sow pools, multi-site farms and farms producing breeding animals. These classifications are relevant because of the variation in their risk potential. For example, it has been hypothesised that a breeding herd would cause a larger amount of additional infections than other farm type (Terpstra, 1988; Rosengren et al., 2002). These different types of farm may be used to predict the risk they cause to other farms and consequently to focus surveillance and monitoring to the sub-populations that induce the highest risk for the production chain. In principle, this type of classification is applicable if a farm type defines precisely the behaviour of farms. However, this is not always the case, since farms may perform different combinations of production.

* Tapani Lyytikäinen, Finnish food safety authority (Evira), Risk assessment unit, Mustialankatu 3, 00790 Helsinki, Finland. E-mail: tapani.lyytikainen@evira.fi

Herd-size may also be important for directing the surveillance effort between farms (Rautiainen et al., 2001; Alban et al., 2002). In Denmark, for instance, herds producing less than 200 slaughter pigs per year are not a part of Salmonella surveillance in the new surveillance program introduced in 2001 (Alban et al., 2002). Farms of small size (measured as number of pigs) have been shown to pose a lower risk to other farms than larger farms (Elbers et al., 2001; Crauwels et al., 2003). One of the assumptions may be that bigger farms have more contacts with other farms and thus pose a higher risk in general. Although this might be true in general, exceptions may induce an unpredictable risk and therefore targeted surveillance based only on farm size may be hazardous.

Another employed criterion for targeting surveillance is the value of animals or operations. For example, only animals in artificial insemination stations (all individuals), performance test stations (individuals to be slaughtered) and elite breeding herds (a sample) are sampled for PRRS in Finland. In addition to these herds, the housed wild boars are also sampled for classical swine fever and swine vesicular disease (Tammiranta & Maijala, 2005). The risk-classification may identify the high risk farms and then it may be possible to improve the biosecurity level of these farms and to modify their operations to decrease the risks.

With endemic diseases, recognising the risk factors and targeting the disease control efforts may be based on past experience and the prevalence of the infection in the country. Specialists, who have experience with the specific infectious disease and with the production chain, are also important. Endemic diseases are often intensively studied and therefore, the risk factors leading to infections and the methods to detect the diseased farms are usually well understood. Experiences with exotic or (re-)emerging infectious diseases are, however, limited or absent, for example, in Finland. Hence, monitoring and surveillance planning have to be undertaken by other methods than for endemic disease.

Simulation provides one option for studying the dynamics of infectious diseases. Simulation models have been applied for instance to study the factors that influence the size and the duration of classical swine fever (CSF) epidemics in the Netherlands (Mangen et al., 2002). Also, the foot and mouth disease (FMD) epidemic that occurred in 2001 in the UK was mimicked afterwards by simulations (Stevenson, 2003). The outcomes of simulations are easily interpreted because they produce numerical estimates for the relevant parameters of the epidemic. A reasonably constructed model can also be used to study the spatial and temporal complexities of epidemics, as shown in several earlier simulation studies on CSF and FMD (e.g. Mangen et al., 2002; Mourits et al., 2002; Velthuis & Mourits, 2007). When a simulation model takes into account farm specific information, which is relevant for the spread of the disease, the influence of different farm specific factors on the expected epidemics may be studied.

The purpose of this study was first to simulate the spread of an infectious disease, namely FMD, among Finnish pig farms. The second aim was to use the outputs from the simulation model, specifically the probability of an epidemic occurring, and the magnitude and duration of epidemics, to classify the Finnish pig farms. In other words, the farms were risk-classified on the basis of their potential to further spread the infectious disease. Second, the different characteristics of farms, for example herd size and number of neighbours, were used to predict the risk-classes. Here, the aim was to investigate if the characteristics of farms predict the risk-classes (i.e. the risk they cause to other farms). This is the first time, to our knowledge, that simulation model results have been used to classify all pig farms in the country with regards to the risk they induce to other farms.

MATERIALS AND METHODS

The Finnish simulation model was developed to study the spread of FMD among pig farms in Finland and is based on an earlier model applied to CSF risk assessment in Finland (Raulo & Lyytikäinen, 2005; Niemi et al., 2005). FMD is regarded as a good 'model disease' because it is highly contagious, transmitted via several routes and literature on the disease is extensive (Donaldson & Alexandersen, 2002; Stevenson, 2003, Sutmoller et al., 2003; Taylor et al., 2004; Grubman & Baxt, 2004). Therefore, FMD simulation provides a model for the worst case scenario on the spread of diseases, in general. The Finnish model is similar to several other simulation models that have been applied to study the relationship between FMD spread and production structures or control policies. Models such as Interspread plus, InterFMD and NAASDM (Sanson, 1993; Mourits et al., 2002; NAASDM Project team, 2007; Schley, 2007; Velthuis & Mourits, 2007) share similar features with the Finnish FMD model. These models all use the concept of transmission probabilities and contacts to estimate the number of infections. In addition, farm locations and farm types are used to parameterise models although the parameterisation itself may vary. They can be thought of as micro simulation models because they try to simulate epidemics between individual farms (Schley, 2007).

Simulation model

The Monte Carlo simulation model was developed in the Matlab 7.3 (Mathworks Inc., MA, USA) environment. The Econometrics toolbox (Le Sage, 2002) was used for generation of random variables. The model consists of both spatial and temporal structures that are in part defined by stochastic features. Events were simulated on a daily basis, with the first simulation day representing the infection day of the primary farm.

The simulation was started separately from each of the Finnish pig farms. The starting farm in each iteration is defined as the primary farm. The source of the infection in the primary farm is assumed to be unknown. In order to create an epidemic, the primary farm must infect at least one other farm.. The infected farms can then induce further infections during their infective period. The infective period starts 7 days after the infectious contact and continues for 3 weeks without restrictive measures (see below for restrictive measures). Transmission is possible when the susceptible farm receives at least one infective contact. Five categories of contact are included in the model: receiving living pigs (high risk contact), a visit by a livestock transportation vehicle or person visiting animal holdings (medium risk contact), a person visiting at the farm (low risk contact), having an infected farm within 1.5 km, and having an infected farm within 1.5 and 3 km. The two last contact types include the so-called neighbourhood transmission, where the vector is not known (Annex 1).

The contacts have different transmission probabilities which define the probability of a contact initiating an infection (Annex 1). To estimate the number of animal and vehicle contacts and to selecting the contact farms, animal movement databases (see below) were read within the infective period of a farm. Neighbourhood spread contacts were based on the farm locations described as coordinates, with distances being calculated using Pythagora´s theorem. Other contact types were estimated according to the parameters given in Annex 1.

In the model, it is assumed that the first detection takes place on the primary farm and by the farmer. Thereafter, detection of infection on a farm may occur in several ways. An infected farm could be detected by farmers, contact tracing or by routine screening (clinical and /or serological) in restriction zones (a protection zone of 3 km, and a surveillance zone between 3

and 10 km) around the infected farms. In detected and suspected farms, pigs are culled and the animal holdings are cleaned and disinfected. For most types of contacts, the infective period ends on the day when restrictive measures are lifted. The infective period of neighbourhood spread of a farm ends when the farm is initially cleaned. For the farms in the restriction zones, the contacts are limited and the farmers are obliged to inform the officials of any signs of disease. The pig farms in the protection zones are visited by a screening team within one week. For the protection zones, serological screenings are performed after 30 days, and for the surveillance zones clinical screenings are performed 20 days, after the last confirmed infected premise on the zone has been initially cleaned. The restrictive measures are applied to the traced contact farms. These farms are visited for clinical inspection within one week from tracing. An iteration is terminated when all infected farms are detected and cleaned.

Data for the simulation model

The Finnish pig farm registry, maintained by authorities, was used as the source data for the number of piglets, sows, hogs and finishing pigs on the farm and the locations of the farms in the year 2006. The registry was assumed to cover the majority of the Finnish pig farm population. The farms that did not keep pigs and did not sell or buy pigs in 2006 were excluded from our database. Hence the number of farms in the model was 3229, each of which had on average 552 pigs.

The classification of Finnish pig farms was based on the type of pigs they managed. In 2006, there were 1092 sow herds, that delivered mostly piglets and lesser degree finishers for slaughter, 1090 finisher herds that delivered pigs for slaughter and received piglets from other farms and 1042 mixed herds that perform a combination of the other types of production.

A pig movement database was constructed, based on the official pig movement registry for 2006. The registry contained information on the sold, purchased and transported pigs between farms and sites in Finland. The full database of pig movements between farms contained 34000 notifications. The transportations of pigs for slaughter were retrieved directly from the pig movement database and contained over 60000 notifications of farmers that delivered pigs to slaughterhouses.

The frequencies of person visiting pig farms or production units were estimated from a questionnaire study, performed during the spring of 2007. There were 525 respondents altogether, corresponding to a response rate of 47%. Prior to the analysis, the farms that responded were divided into three main production types (see above): Sow farms (I_s Annex 1), mixed farms (I_m, Annex 1) and finisher farms. The number of pigs on the farm was standardised within each production type. The parameterisation of frequencies was performed using generalized linear models (McCullaugh & Nelder, 1989) in the SPSS statistical package (15.01, Illinois, USA). Developed functions are given in Annex 1.

Simulations

The simulations were started on three dates: 01/01/06 (day 0), and 90 and 180 days later. On each date, 600 000 iterations were performed, which means that every farm acted as the primary farm in approximately 186 iterations per starting date. Because the primary farms were selected randomly, the number of iterations per primary farm varied within a date and also between the starting dates.

The output of the simulations consisted of the number of infected farms at the end of each epidemic, the duration of the epidemic, the identity of infected farms and the identity of the farm which was the source of infection.

Outcomes of the simulations and risk-classification

The risk-classification was based on the simulated results of four variables (1-4): 1) The probability of further spread from each primary farm was calculated: the number of iterations, which caused further spread, was divided by the total number of iterations that were started from the farm. Similarly, 2) the mean and 3) maximum number of subsequently infected farms during an epidemic and 4) the duration of epidemic were estimated for each farm according to those iterations where it was the primary farm.

Farms were classified into four risk-classes (very high, high, medium and low) based on the four variables (see above) using the K-means clustering method. Because there were differences in the variances of variables, they were standardized before clustering by subtracting their mean and then dividing the values with standard deviation. K-means clustering started with constructing the initial cluster centres according to the simulation results. After obtaining initial cluster centres, the procedure assigned farms (according the simulation results) to clusters based on distance from the cluster centres. The locations of cluster centres were then updated on the basis of mean values of simulation results in each cluster. These steps are iterated until any reassignment of simulated results of farms would make the clusters more internally variable or externally similar.

Development of the predictive model

To predict the risk category of a farm, based on its characteristics (see below), an ordinal regression procedure using the logit link function was carried out in SPSS. The probability of farms belonging to a category can be estimated by applying Eq.(1) sequentially:

$$P_{low} = 1-[(e^{B-B_{low})} / (1+e^{B-B_{low}})]$$

$$P_{medium} = 1-[(e^{B-B_{medium}}) / (1+e^{B-B_{medium}})]-P_{low}$$

$$P_{high} = 1-[(e^{B-B_{high}}) / (1+e^{B-B_{high}})]-[P_{low}+P_{medium}]$$

$$P_{very\ high} = 1-[P_{low}+P_{medium}+P_{high}] \quad (1)$$

In Eq. (1) P_{low}, P_{medium}, P_{high}, $P_{very\ high}$ = the probability of a farm belonging to the class given by the subscript, B_{low}, B_{medium} and B_{high} = threshold values for the low, medium and high classes. The predicted class was the one with the highest probability. The initial set of explanatory variables in the model (i.e. the characteristics of the farm) were: 1) the number of animals on the farm, 2) the number of deliveries from the farm to other pig farms, 3) the number of batches sent to slaughter during 2006, 4) the number of received pigs/transports during 2006, 5) the number of pig farms within 1.5 km and 6) the number of pig farms between 1.5 and 3 km from the farm.

The variables were included in the model if the Wald-test statistic indicated that the probability of the parameter being zero was < 0.05.

The sensitivity and the specificity of the predictions for each class were estimated by formation of a confusion matrix. In other words, by cross tabulating the farm classification by clustering with the predicted classes of farms based on the ordinal regression models. The overall statistical significance of the models was tested against the null model using the likelihood –ratio test. Analyses were performed by in SPSS 15.01 statistical package (SPSS Inc., IL, USA).

RESULTS

Results from the simulations

The results (farm specific values) were estimated for each farm based on the outputs of the iterations for which it was the first infected farm, i.e. the primary farm and are shown in Table 1. Here, the farm specific mean was derived from the separate iterations which started from each of the 3229 farms. The range describes the minimum and maximum of the farm specific means. The mean of all farms is calculated from the farm specific means. Spearman correlations were calculated for farm specific values between different simulation days (day 0, 90 and 180). There were large differences between individual farms. In particular, the farm specific mean number of subsequent infected farms varied from 1 to 19 infected farms. The range of subsequent infected farms also varied between starting days: the maximum mean for all farms was almost twice as high on day 0 as it was on day 180 (Table 1). The number of infected farms, when the mean of all farms was considered, differed slightly between the starting dates but the relative differences between dates were rather small (under 5%).

There were differences between individual farms within a starting date with regards to the probability of further spread. There were farms which rarely initiated further spread, but also farms which infected other farms each time they were the primary farm. The average for all farms for the probability of further spread did not vary between starting days (Table 1). The average length of epidemics appeared to be relatively stable since both the day-to-day and the range of farm specific means were relatively small (Table 1).

The strongest correlations between the simulation dates were in the farm specific probabilities of further spread. Also, the number of infected farms and the duration of the epidemic resulted in relatively high correlations between the starting dates. The lowest correlations were in the maximum number of infected farms between different starting dates (Table 1).

Table 1. Simulation results from the Finnish FMD model

VARIABLE	SIMULATION DAY 0	90	180	COMBINED	RANGE OF CORRELATIONS
Mean number of infected farms[a]					
mean of all farms	3.50	3.36	3.44	3.47	
range of farm specific means	1-27	1-21	1-16	1-19	0.792-0.832
Maximum number of infected farms					
range of farm specific maximum	1-209	1-91	1-112	1-209	0.577-0.654
Probability of further spread					
mean of all farms	0.62	0.62	0.62	0.62	
range of farm specific means	0-1	0-1	0-1	0-1	0.931-0.941
Duration of epidemic (days)					
mean of all farms	38	38	38	38	
range of farm specific means	31-60	31-64	31-60	30-60	0.863-0.885

[a] simulations with no spread are excluded prior to calculations

The mean and maximum number of infected farms and the probability and the duration of epidemics caused by a specific primary farm depended on each other. The probability of further spread appeared to have a positive non-linear relationship with the mean number of infected farms (Fig. 1a). The maximum number of infected farms appeared to have a positive relationship with the mean number of infected farms (Fig. 1b). The duration of the epidemic was almost linearly related with the mean number of infected farms and had the highest positive correlation with the number of infected farms (Fig. 1c).

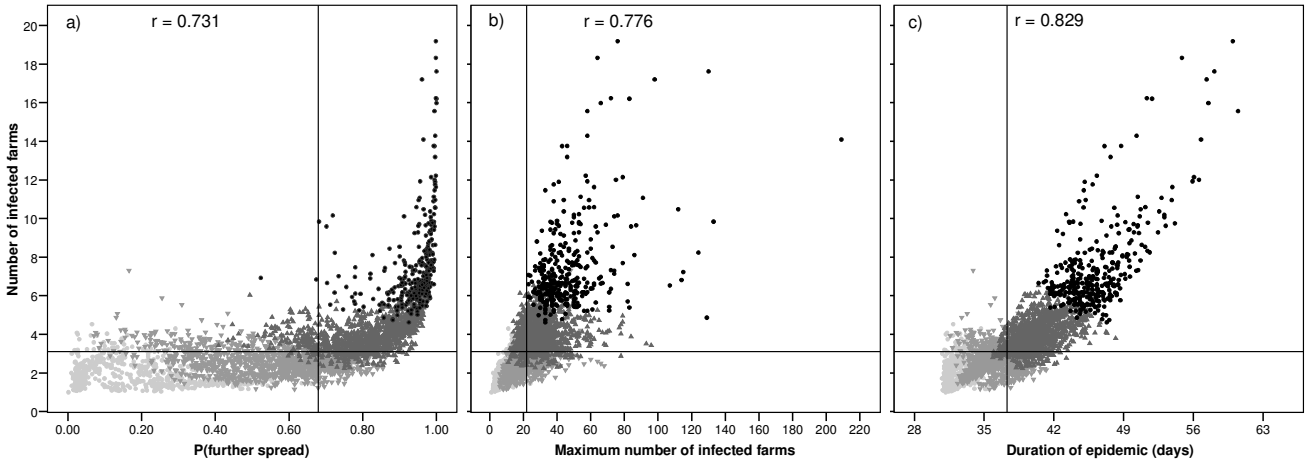


Fig. 1. The relationship between the mean number of infected farms and a) the probability of further spread, b) the maximum number of infected farms c) the duration of epidemics (days). Each point represents the mean or maximum value of the farm estimated from the iterations when the farm was primary infected farm in the country. The risk-classes are represented by the color and shape of the points: black dot = very high, dark grey triangle = high, grey reverse triangle = medium, light grey dot = low. Solid lines represent median for the variables. Spearman correlation coefficients are given

Risk-classification

Pig farms were divided into four risk-classes based on the results of the iterations which started from the particular farm. The classification was based on the probability of an epidemic occurring and the magnitude and duration of epidemics which did occur (Fig. 1a - c). The highest class promoted, on average, the largest and longest epidemics with the highest probability of further spread. In addition, their potential to produce extremes was also the highest. Classes deviated as the differences of class centres appeared to be at least one standard deviation apart (Table 2). Thus, the farms were separated in different classes by the mean of all four variables that were used in clustering. The highest class constituted approximately 10% of the Finnish pig farm population, while the two intermediate classes presented the majority, (>70 %) of the population.

Table 2. The four risk-classes of Finnish pig farms (n=3229) based on their ability to promote spread of FMD. Values represent means (standard deviation).

VARIABLE	RISK-CLASS			
	low	medium	high	very high
Probability of further spread	0.19 (0.15)	0.56 (0.16)	0.81 (0.10)	0.94 (0.07)
Mean number of infected farms	1.94 (0.65)	2.60 (0.64)	3.97 (0.78)	7.48 (2.30)
Max. number of infected farms	9.00 (4.29)	19.59 (7.64)	29.89 (10.83)	47.37 (19.99)
Duration of epidemic (days)	32.30 (0.98)	35.77 (1.66)	40.22 (2.06)	46.34 (3.29)
% of farms in Finland	17.8 %	36.8 %	35.1 %	10.3 %

Table 3 gives the ratio of the simulated outcomes (% of infections that farms in the receiver risk-class comprise from all the infections induced by the risk-class of source farms) to the expected outcomes (% of farms within the risk-class in Finland). Values can be interpreted as percentages within risk-class of source farm by multiplying the ratios by the expected value (given on the bottom row). The members of the very high risk-class tended to induce infection in other members of the very high risk-class. In contrast, they infected the members of the low and the medium risk-classes less often than expected. The members of the medium and the high risk-classes infected members of the high risk-class most frequently. It is notable that all risk-classes tended to infect the members of high and very high risk-classes more often than would be expected by the proportion of farms in the high and very high risk-classes in the Finnish pig farm population (Table 3).

Table 3. The ratio of the simulated outcomes to the expected outcomes. Values above 1 show positive tendency from source to receiver risk-class.

RISK-CLASS OF THE SOURCE FARM OF AN INFECTION	RISK-CLASS OF THE RECEIVER OF AN INFECTION			
	low	medium	high	very high
low	0.55	1.11	1.03	1.26
medium	0.19	0.82	1.38	1.78
high	0.06	0.45	1.49	2.90
very high	0.03	0.22	1.04	5.27
expected % of infections	17.8	36.8	35.1	10.3

Predicting the risk-classes, based on the simulation results (see above), using the farm specific characteristics as predictors, was undertaken using ordinal regression models (logit link function). The statistically significant predictors were the number of pig deliveries to other farms during within a year (D), the number of pig deliveries to slaughter (S), the number of farms within 1.5 km from the farm in focus (N_{15}), and the number of pig farms within 1.5-3 km from the farm (N_{30}). The equation of the model and the parameter estimates (± SEM given in parenthesis) for the characteristics are shown in Eq.(2). B is the predictive part of regression function.

$$B=0.128(\pm0.004)D+0.086(\pm0.003)S+0.714(\pm0.034)N_{15}+0.775(\pm0.033)N_{30} \quad (2)$$

The estimates for the threshold values given in Eq.(1) are

$$B_{low} = 1.266$$

$$B_{medium} = 4.646$$

$$B_{high} = 9.402$$

This model was significantly better than null model (Chi-square = 3446.9, df = 4) and it predicted risk-classes of the farms quite well, (Nagelkerke´s pseudo R^2-value of the model was 0.711). The sensitivity of the predictions varied from 57.8 to 76.3%, being the lowest in the highest risk-class (Table 4). The specificity of classification was usually higher varying from 64.7 to 77.4%. The farms predicted to be in the highest class belonged to the highest class by clustering on 68.2% of occasions (Table 4, Fig. 2a).

Table 4. The confusion matrix for Finnish pig farms. Values are given as percentages of farms categorized by simulation falling into given predicted risk-class.

PREDICTED	SIMULATED RISK-CLASS			
RISK-CLASS (N)	low	medium	high	very high
low (456)	61.4	8.3	0.3	0.3
medium (1401)	38.1	76.3	23.3	3.6
high (1089)	0.5	15.2	68.6	38.3
very high (283)	0	0.1	7.9	57.8
simulated N	575	1187	1133	334

In addition to the explanatory variables mentioned above, a simpler predictive model was employed and evaluated. In this case, only the number of pigs on the farm and the production type of the farm were used as explanatory variables. The model had clearly lower predictive value than the previous model (Fig. 2a - b). Especially poor regression model results were achieved on identifying the members of highest risk-class, where only 7.2% were identified correctly. The sensitivities were higher for the other classes (57.4 - 81.6%). The specificity of the predictions were also lower (34.8 - 44.2%) than by the model of Eq. (2). Generally, this model predicted poorly the risk-classes of the farms (Nagelkerke´s pseudo R2-value of the model was 0.172).

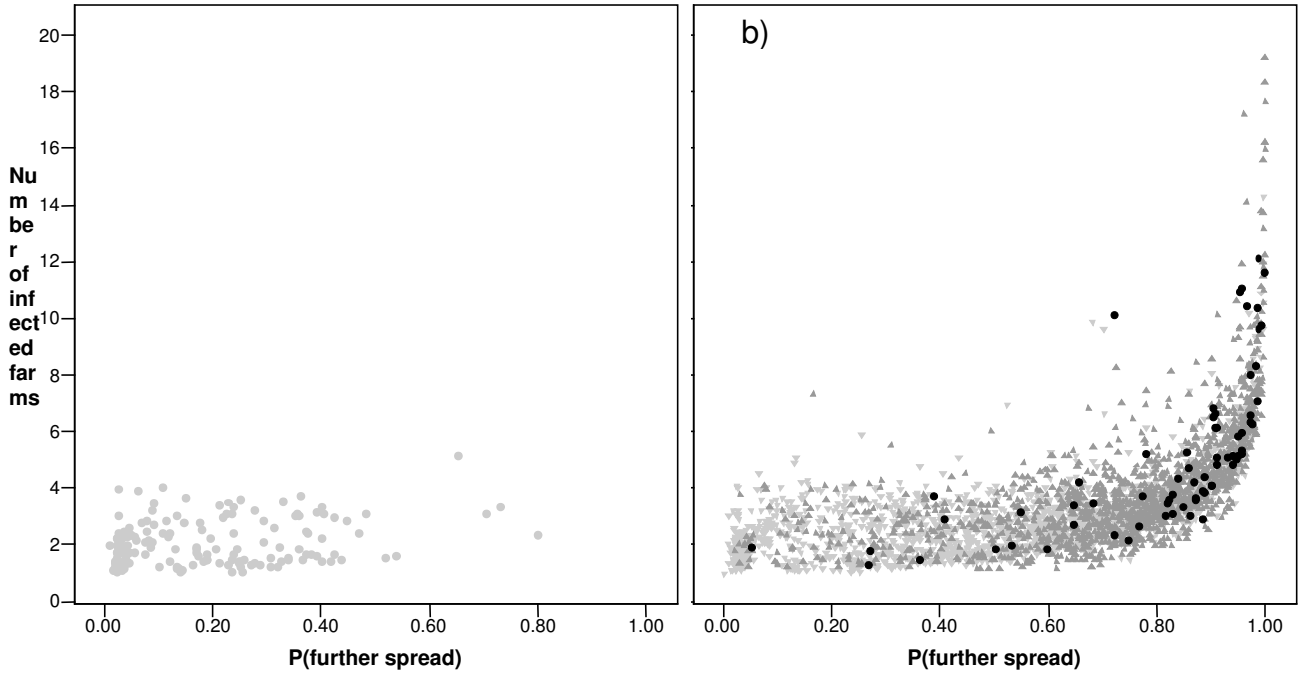


Fig. 2. Predicted risk-classes generated from the simulation outcomes (Fig. 1) of Finnish pig farms by ordinal regression a) Explanatory variables were the exact number of neighbouring farms and the number of pig deliveries to other farms and to slaughter, see Eq.(2) b) Explanatory

variables were the farm type (finisher, mixed or sow farm) and the number of pigs on the farm. Predicted classes are represented by shape and colour of the points: black dots = very high, dark grey triangles = high, grey reverse triangles = medium, light grey dot = low.

DISCUSSION

The Finnish FMD simulation model was used to simulate the spread of FMD when the epidemic starts from a primary farm. There were differences between farms in their potential to further spread infection. When transmission from the primary farm took place, 3.5 other farms became infected on average and the mean duration of an epidemic was 5 weeks. These results are in line with earlier real experiences with FMD (Sutmoller et al., 2003; McClaws & Ribble, 2007) and are similar to the simulation results in a sparsely populated area in the Netherlands by Velthuis & Mourits (2007).

Finnish pig farms appeared to present differences in their ability to initiate further spread of infectious diseases. However, the three starting dates gave very similar results for each of the farms, which indicates that there are true differences between farms instead of only random variation or day-to-day fluctuations. The mean and maximum numbers of infected farms caused by a primary farm varied between the starting dates of the simulations. However, this was predictable since the simulation results were directly dependent on the amount of contacts of the primary farm and on the farms connected with primary farm. The durations of a subsequent epidemic also varied and were dependent on the official control measures and to a lesser extent, the size of epidemic and thus the identity of primary farm. In contrast, the most stable measure, i.e. the probability of further spread, depended only on the contact network of the primary farm during the infective period. There were primary farms which did not transmit the infection further in any of the simulations, whereas some of the farms always caused an epidemic.

The differences between the simulation results for each farm were prerequisites for the risk-classification. Because the simulation results correlated, they produced four consistent clusters. The farms of the highest spreading risk-class produced the worst expectations in every aspect of epidemic. We did not elaborate our attempts to define the most valid clustering method. In addition, we do not make any statements on the general number of relevant subgroups and actually more or less subclasses may be more sensible for simulation or practical reasons.

Two types of predictive models were developed for the prediction of risk-classes The idea was that if the farm level characteristics would predict the risk-class of a farm well, they could be used for the risk-classification without simulation modelling or new simulations. Hence, a new farm, for instance, could be risk-classified based on the farm level characteristics. The predictive model that used the number of animal deliveries to slaughter and to other farms and the number of neighbourhood farms within 3 km of a farm classified the farms similarly to how they were classified based on the simulation results (which were used as gold standard). Thus, they may offer a way to direct surveillance and monitoring towards the highest risk-classes because the separation of the low and the very high risk-classes was also almost complete. However, they failed to reach a level of sensitivity and specificity, that would mean that they could be used as a definitive rule for risk-classification. Hence, using predictive models for risk-classification, and consequently targeting sampling effort should be undertaken with an awareness of the limits of the predictions: they are indicative not definitive.

Predicting the risk-classes using the size and production type of a farm as the explanatory variables had poor predictive value. This might indicate that although these variables correlated with the relevant operations of the farm, they failed to capture the factors that promote spread. Improvement of the predictive value by developing more elaborate production type classification may be valuable but would require that the farm registry contains similar typing information. Risk-classification by identification of relevant farm types should be studied in more depth since it may have high practical value.

Members of the very high risk-class tended to infect other members of the very high risk-class. On the other hand, one key component of the members of very high risk-class seemed to be that the members of the very high risk-class infected members of other risk-classes to a lesser extent than would be expected. Thus, results indicate that an integral part of the definition of the very high risk-class concerns their interconnections. From a practical point of view it is of concern that members of the very high risk-class also receive infections from the members of the other risk-classes and thus the involvement of the very high risk-class is one of the main key factors in large epidemics. In the future, this could be partly prevented by directing the large production units away from the existing pig production units. Decreasing animal movements risks could be achieved for instance by limitations to the number of trading partners. High risk farms should also be the focus when infectious diseases are surveyed and monitored.

Because heterogeneity of networks has been shown to influence the dynamics of epidemics, it should be included in any epidemic model (Barthélemy et al., 2004; Barthélemy et al., 2005, Shirley & Rushton 2005; Crépey et al. 2006; Colizza et al., 2006; Bigras-Poulin et al., 2007) if the aim is to apply outcomes for risk-classification of farms. The Finnish FMD model uses an animal movement database and historic nonparametric information about the animal movement network throughout the country. This approach replicated animal transportations between farms exactly as they happened in 2006, thus containing a large part of the true operational network of a farm. The model also allows also different unidentified factors to influence the outcome and might lead to identification of new relevant characteristics or factors. The registers which were used to construct the animal movement database were also applied in official statistics of Finland and thus can be regarded as reliable. The Finnish model is able to produce a very large number of iterations which was a prerequisite for successful risk-classification of Finnish pig farms.

A relevant aspect of the classification of farms is the relative differences between transmission probabilities. The applied transmission probabilities were based mainly on the work of Stevenson (2003). An earlier study has shown that the transmission probabilities applied in our model produced fairly good agreement with the UK 2001 FMD epidemic simulated by Interspread (Stevenson 2003) – thus we are confident of their applicability for the Finnish FMD model. Even if absolute values are not exact, our simulations may be applicable for classification purposes if the relative differences between the transmission probabilities are correct.

Much of an epidemic was dependent on the identity of primary farm. This may be a result of the small overall size of the epidemics, due to which the identity of the primary farm had a relatively high importance. In Finland, the pig farm density is low since there are only 87 pigs / km2 on the densest part of the country. This is much lower than, for instance, in The Netherlands, where even sparsely populated areas may have more pigs and farms (Mangen et al., 2002; Mourits et al., 2002; Velthuis & Mourits, 2007). Therefore, similar kinds of simulations should be performed in countries which have intensive production structures or higher farm

densities and may produce large epidemics to evaluate the importance of the identity or characteristics of a primary farm for the epidemic in general.

Finland's pig farm population consisted of 3229 farms in 2006, which was small enough to simulate the spread of a FMD starting from every farm in the country. If a similar approach could be used for larger groups of farms it would enable the estimation of risk that pig farms may pose to sheep and cattle farms. The inclusion of cattle and sheep farms into the simulation would improve the general predictive value of the model especially in the lower risk-classes where this interspecies risk might elevate the risk-class of the farm but the approach would also require modifications for the model.

More studies are needed to certify usefulness and limitations of risk-classification of individual farms by simulations. However, the results indicated that the simulation based classification could be a valuable method for practical risk management and may be used in the future for instance in targeting of sampling efforts.

ACKNOWLEDGEMENTS

The research is partially funded by National agricultural development fund (MAKERA). We would like to also thank Saara Raulo and Jarkko Niemi for their effort in the planning of the original classical swine fever model. We are grateful for the critical comments of Leena Sahlström, Kimmo Suominen and Sonja Koski.

REFERENCES

Alban, L., Stege, H. and Dahl, J. (2002). The new classification system for slaughter-pig herds in the Danish Salmonella surveillance-and-control program. Prev. Vet. Med. 53, 133-146.

Barthélemy, M., Barrat, A., Pastor-Satorras, R. and Vespignani, A. (2004). Velocity and hierarchical spread of epidemic outbreaks in scale-free networks. Phys. Rev. Lett. 92, 1-4.

Barthélemy, M., Barrat, A., Pastor-Satorras, R. and Vespignani, A. (2005). Dynamical patterns of epidemic outbreaks in complex heterogeneous networks. J. Theor. Biol. 235, 275-288.

Bigras-Poulin, M., Barford, K., Mortensen, S. and Greiner, M. (2007). Relationship of trade patterns of the Danisg swine industry animal movements network to potential disease spread. Prev. Vet. Med. 80, 143-165.

Colizza, V., Barrat, A., Bathelemy, M. and Vespignani, A. (2006). The role of the airline transportation network in the prediction and predictability of global epidemics. PNAS 103, 2015-2020.

Crepéy, P., Alvarez, F.P. and Barthelemy, M. (2006). Epidemic variability in complex networks. Phys. Rev. E 73, 1-7.

Crauwels, A.P.P., Nielen, M., Elbers, A.R.W., Stegeman, J.A. and Tielen, M.J.M. (2003). Neighbourhood infections of classical swine fever during the 1997-1998 epidemic in the Netherlands. Prev. Vet. Med. 61, 263-277.

Donaldson, A.I. and Alexandersen, S. (2002). Predicting the spread of foot and mouth disease by airborne virus Rev. Sci. Tech. Off. Int. Epiz. 21, 569-575.

Elbers, A.R.W., Stegeman, J.A. and de Jong, M.C.M. (2001). Factors associated with the introduction of classical swine fever virus into pig herds in the central area of the 1997/98 epidemic in the Netherlands. Vet. Rec. 149, 377-382.

Grubman, M.J. and Baxt, B. (2004). Foot-and-Mouth Disease. Clin. Microbiol. Rev. 17, 465–493.

Le Sage, J.P. (2002). Econometrics toolbox. http://www.spatial-econometrics.com/.

Mangen, J-M.J., Nielen, M and Burrell, A.M. (2002). Simulated effect of pig-population density on epidemic size and choice of control strategy for classical swine fever epidemics in The Netherlands. Prev. Vet. Med. 56, 141-163.

McLaws, M. and Ribble, C. (2007). Description of recent foot and mouth disease outbreaks in nonendemic areas: Exploring the relationship between early detection and epidemic size. Canadian Veterinary Journal 48, 1051-1062.

Mintiens, K, Laevens, H, Dewulf, J, Boelaert, F, Verloo, D and Koenen, F. (2003). Risk analysis of the spread of classical swine fever virus through "neighbourhood infections" for different regions in Belgium. Prev. Vet. Med. 60, 27-36.

Mourits, M.C.M., Nielen, M. and Léon, C.D. (2002). Effect of control measures on the course of simulated foot and mouth disease epidemics started on different farm types in various Dutch areas. *In* Proceedings of Society for Veterinary Epidemiology and Preventive Medicine, Cambrigde, UK, pp.190-200.

McCullagh, P. and Nelder, J.A. (1989). Generalized linear models – Second Edition. Monographs on Statistics and Applied Probability 37. Chapman & Hall, Boca Raton, Florida, USA. 511p.

NAADSM- Project-team (2007). North American Animal disease spread model, free program distributed via the internet at http://www.naadsm.org.

Niemi, J.K., Lehtonen, H., Pietola, K., Raulo, S. and Lyytikäinen, T. (2005). Economic implications of classical swine fever in Finland (In Finnish), Maa- ja elintarviketalous 74, 82p.

Raulo, S. and Lyytikäinen, T. (2005). Epidemic outbreak of classical swine fever in Finland, quantitative risk assessment (In Finnish). National Veterinary and Food Research Institute EELA Publications. 06/2005. 192 p.

Rautiainen, E., Konradsson, K., Lium, B., Mortensen, S. and Wallgren, P. (2001). Disease surveillance strategies in swine. Acta. Vet. Scand. 94, 31-42.

Rosengren, H., Rautiainen, E., Lyytikäinen, T. and Maijala, R. (2002). Qualitative risk assessment of the spread of classical swine fever into and within Finland. National veterinary and food research institute EELA publications. 94p.

Sanson, R.L. (1993). The development of decision support system for an animal disease emergency. Ph.D. Thesis, Massey university, Palmerston North, New Zealand, 264 p.

Schley, D. (2007). Recommendations for enhancing confidence in epidemiological models of Foot-and-mouth disease. 37th general session of the European commission for the control of foot-and-mouth disease. 1-17.

Shirley, M.D.F and Rushton, S.P. (2005). Where disease and networks collide: lessons to be learnt from a study of the 2001 foot-and-mouth disease epidemic. Epid. Infect. 133, 1023-1032.

Stevenson M.A. (2003). The spatio-temporal epidemiology of bovine spongiform encephalopathy and foot-and-mount disease in Great Britain. Ph.D. Thesis. Institute of Veterinary, Animal and Biomedical Sciences, Massey University, Pelmmerston, New Zealand. 249p.

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

Sutmoller, P., Bartelling, S.S., Olascoaga, R.S. and Sumption, K.J. (2003). Control and eradication of foot-and-mouth disease. Virus Res. 91, 101-144.

Tammiranta, N. and Maijala, R. (2005). National surveillance programmes for contagious animal diseases in National Veterinary and Food Research Institute EELA, National Veterinary and Food Research Institute EELA Publications 07/2005. 95p.

Taylor, N.M., Honhold, N., Paterson, A.D. and Mansley, L.M. (2004). Risk of foot-and-mouth disease associated with proximity in space and time to infected premises and the implications for control policy during the 2001 epidemic in Cumbria. Vet. Rec. 154, 617-626.

Terpstra, C. (1988). Epizootiology of hog cholera *In:* Classical swine fever and related viral infections, ed. Liess, B., Martinus Nijhoff Publishing, pp. 201-233.

Velthuis, A.G.J. and Mourits, M.C.M. (2007). Effectiveness of movement prevention regulations to reduce the spread of foot-and-mouth disease in the Netherlands. Prev. Vet. Med. 82, 262-281.

Annex 1. Contacts between pig farms, selection of receivers and the transmission probabilities used in the Finnish FMD simulation model over a daily time step.

INCIDENT LEADING TO TRANSMISSION OF INFECTION	FREQUENCY OF POTENTIAL CONTACTS	POTENTIAL FOR CONNECTION	NUMBER OF FARMS CONNECTING	SELECTION OF RECEIVING FARMS	TRANSMISSION PROBABILITY
pig movement					
transportation of live pigs between farms	pig-movement database[a]	1 = contacts, 0 = no contacts[a]	pig movement database[a]	farms registered to receive pigs from the infected source farm [I]	0.4[h]
an unclean transportation vehicle					
transporting pigs between farms transporting pigs to slaughter	pig movement database[a]	1 = contacts, 0 = no contacts[a]	pig movement database[a] ~NEGBIN(S=4, P=0.5)[b]	other farms registered to receive or send pigs by same vehicle as infected farm [I] farms that deliver pigs to the same slaughterhouse as infected farm [I]	0.15[h]
a person					
visiting the production unit	~PO(Eq.(3))[f]	0-0.5	~NEGBIN(S=1, P~UD (min=0.5, max=1.0))	farms from the same operational sectors as infected farm: 86 sectors for veterinarians and 15 sectors for production consultants; 5 sectors for AI-technicians; 132 sectors for substitutes traced contact farms 86 sectors for culling teams	0.15[h]
visiting without entering to production units	~PO(Eq.(4)) [f]		~NEGBIN(S=1, P~UD (min=0.5, max=1.0))		0.005[h]
working as substitute	0.0124[g]	~BETA(21, 110)[c]	2		0.15[h]
screening for FMD	1 for screening day of undetected infected farm	~UD(min=0, max=1.0)	~UD(min=0, max=2)		0.005[h]
culling pigs	4 for culling day	~UD(min=0.01, max=0.1)	1-4[e]		0.005[h]
neighbourhood spread					
neighbouring pig-farms within 1.5 km	once a day during infective period	1 = farms in range 0 = no farms in range[d]	number of farms in range[d]	all farms within range[d]	0.009[i]
neighbouring pig-farms between 1.5 to 3 km					0.0035[i]

[a]pig movement registry, 2006; [b]Rosengren et al., 2002; [c]Raulo & Lyytikäinen, 2005; [d]pig farm registry, 2006; [e]Jarkko Niemi, unpublished; [f]questionnaire 2007; [g]Statistics of social insurance institution of Finnish farmers; [h]Stevenson, 2003; [i]Taylor et al., 2004. NEGBIN = Negative binomial distribution, UD = uniform distribution, BETA = Beta distribution, PO = Poisson distribution NO = Normal distribution. Eq. (3): lambda = $e^{(0.381I_s+0.529I_m+2.224)}$ Eq. (4): lambda = $e^{(-0.413I_s+0.344Z_n+3.596)} \cdot NO(Mean=0.833,Std=0.386)$, in equations: I_s = Indicator of sow herd, I_m = Indicator of mixed herd, Z_n=standardized number of pigs on the farm.

NO INCREASED RISK WITH SHORTER WAITING PERIOD AFTER FOOT-AND-MOUTH DISEASE VACCINATION

C.J. DE VOS*, M. NIELEN, E. LOPEZ, A.R.W. ELBERS AND A. DEKKER

SUMMARY

Emergency vaccination is the most effective control strategy for foot-and-mouth disease (FMD) epidemics in densely populated livestock areas, but results in a six-month waiting period before exports can be resumed. Due to resulting severe economic consequences for pig exporting countries a one-month waiting period, to be based on negative test results in the final screening, was considered for the European Union. The goal of this study was to analyse the risk of exporting pig carcasses from a vaccinated area (a) directly after final screening and (b) after a six-month waiting period. A risk model was built to calculate the probability that a carcass derived from an FMD-infected pig would be processed (Pcarc). Leading variables were herd prevalence (PH), within-herd prevalence (PA), and the probability of detection at slaughter (PSL). PH and PA were estimated using Bayesian inference. Model calculations indicated that P_{carc} is on average 2.0×10^{-5} directly after final screening and 1.7×10^{-5} after a six-month waiting period. Thus the additional waiting time thus only slightly reduced P_{carc}.

INTRODUCTION

In the European Union (EU), control of foot-and-mouth disease (FMD) has been based on a non-vaccination policy since the early 1990s and epidemics are contained primarily by movement controls and stamping out of infected and contact herds. Modelling studies have, however, indicated that vaccination of susceptible animals is the most effective control strategy if an epidemic occurs in a densely populated livestock area, such as the south-eastern part of the Netherlands (Tomassen et al., 2002).

When emergency vaccination is applied, export from the affected area can only be resumed six months after the last detected outbreak or after the last vaccination, whereas exports can be resumed three months after the last detected outbreak if no vaccinated animals are present (CEC, 2003; OIE, 2007). For this reason, the 2001 epidemic in the Netherlands was controlled by a vaccination-to-cull strategy, i.e., all animals in the affected area were vaccinated and subsequently killed and destroyed (Bouma et al., 2003). Although effective, this approach raised ethical questions about the need for large-scale slaughter of vaccinated, but healthy animals. Because the Netherlands is a major exporter of pigs and pork, a vaccination-to-live policy is, however, only economically attractive if products derived from vaccinated animals would be accepted by its trading partners soon after the end of an epidemic. This will only be the case if the accompanying FMD risk is acceptable.

* Clazien de Vos, Department of Virology, Central Veterinary Institute of Wageningen UR, P.O. Box 65, 8200 AB Lelystad, the Netherlands. E-mail: clazien.devos@wur.nl

Since EU legislation prescribes a final screening 30 days after the last outbreak or vaccination to ascertain freedom of disease, a one-month waiting period could be considered. This final screening should include all farms in the affected area and would consist of clinical inspection of all animals and serological testing of a sample of animals to detect a 5% prevalence with 95% confidence (CEC, 2003). When positive results are obtained, the epidemic would be assumed not to have ended. Given the socio-economic consequences of positive results, all will be done to exclude false positives resulting in an ultimate test specificity of 1. When only negative results are obtained, the area is assumed to be free from FMD, although some infected animals might have been missed due to sampling and use of imperfect tests (test sensitivities < 1).

The goal of this study was to analyse the risk of exporting pig carcasses from a vaccinated area after both a one-month and a six-month waiting period.

MATERIALS AND METHODS

A risk model was developed to calculate the probability that a carcass derived from an FMD-infected pig would be processed (P_{carc}). This probability is calculated as in Eq.(1):

$$P_{carc} = P_H \times P_{Ai} \times \left(1 - P_{SL}\right) \tag{1}$$

where P_H is the herd prevalence giving the probability that a pig originates from an infected herd, P_{Ai} is the animal prevalence in infected herds giving the probability that an individual pig is infected , and P_{SL} is the probability that an infected pig is detected at slaughter.

The model is a stochastic simulation model. Model calculations were performed in Microsoft Excel and @Risk (Palisade Corporation, 2002), running 10,000 iterations for each scenario.

Model calculations

Figure 1 gives a schematic representation of the model calculations. In each iteration, one value is sampled for the probability that an infected pig is detected at slaughter (P_{SL}) and the herd prevalence (P_H). Then, the number of infected herds is calculated by multiplying P_H by the total number of herds in the affected region (N_H). Subsequently, for each infected herd its within-herd prevalence (P_A) and herd size (N_A) are sampled to calculate the number of infected pigs on the farm. The animal prevalence in infected herds (P_{Ai}) in the affected region is then calculated by dividing the total number of infected pigs on all infected farms by the total number of pigs on these farms. Finally, P_{carc} is calculated using Eq.(1).

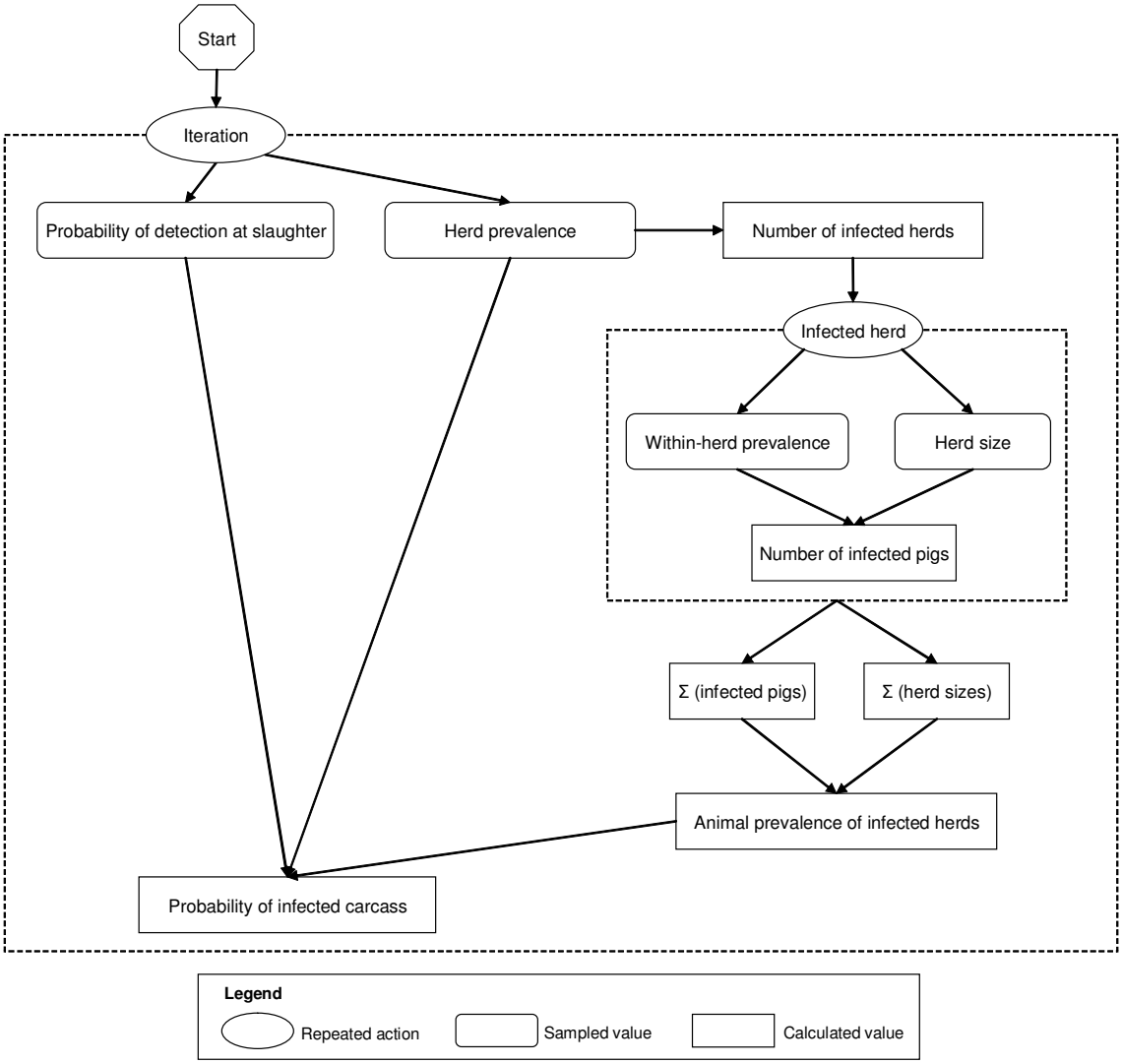


Fig. 1 Schematic representation of the model calculations

Model inputs

An overview of all input parameters in the model is given in Table 1, the majority of which are described in more detail below.

Vaccinated area: Model calculations were performed for a vaccinated area that resembled the densely populated pig area in the south-eastern part of the Netherlands. This area, consisting of the provinces of Noord-Brabant, Gelderland and Limburg, has 6360 pig herds with an average herd size of 1361 pigs (CBS, 2005). Because it is very unlikely that the whole region will be affected during an epidemic, it was assumed that ~5% of the farms, i.e. 318 pig farms, would be vaccinated. This estimate was based on the experience of the 2001 FMD epidemic.

Table 1. Description of input parameters and their values

Input parameter	Abbreviation	Value	Source
Number of vaccinated pig herds in affected area	N_H	318	CBS, 2005
Herd size	N_A	Discrete distribution (16 herd sizes)	CBS, 2005
Probability of detection at slaughter	P_{SL}	Beta(4,12)	Eblé et al., 2004; Eblé et al., 2007; Orsel et al., 2007
Sensitivity of clinical inspection during final screening	Se_c	Beta(27,15)	Eblé et al., 2004; Eblé et al., 2007; Orsel et al., 2007
Sensitivity of serological test (NSP-ELISA)	Se_s	Beta(34,8)	Eblé et al., 2004; Eblé et al., 2007; Orsel et al., 2007
Sensitivity of clinical inspection during waiting period	Se_w	Beta(27,15)	Eblé et al., 2004; Eblé et al., 2007; Orsel et al., 2007
Confidence for sample size calculations	S_c	0.95	CEC, 2003
Prevalence for sample size calculations	S_p	0.05	CEC, 2003
Number of clinical inspections during waiting period	m	5	OIE, 2007
Basic reproduction ratio for within-herd spread	R_0	2.42	Orsel et al., 2007
Herd sensitivity of clinical inspection during final screening	HSe_c	0.884	Calculations
Herd sensitivity of serological test (NSP-ELISA)	HSe_s	0.647	Calculations
Herd sensitivity of clinical inspection during waiting period	HSe_w	0.997	Calculations

Test protocols: Three test protocols were distinguished: clinical inspection during the final screening (*clinical*), serological testing during the final screening (*serology*), and clinical observance in the additional waiting period (*waiting*). *Serology* was based on an ELISA that detects antibodies to non-structural proteins of FMD virus (FMDV). This so-called NSP-ELISA discriminates between infected and non-infected animals regardless of their vaccination status (Brocchi et al., 2006).

Test specificities were set at 1 for all test protocols, because it was assumed that there will be maximum effort to exclude false positives. Test sensitivities of *clinical* (Se_c) and *serology* (Se_s) were based on experimental data (Eblé et al., 2004; Eblé et al., 2007; Orsel et al., 2007). The test sensitivity of *waiting* (Se_w) was assumed to equal Se_c, but was modelled as a monthly measurement because farmers will observe their animals frequently. The test protocol *waiting* thus is a repeated measurement consisting of five additional clinical inspections ($m = 5$) during a six-month waiting period, resulting in a higher probability of detecting infected pigs than with *clinical*.

In the protocol *serology* a sample of pigs is tested. Sample sizes (N_S) were calculated for each herd individually, taking into account Se_s (Martin et al., 1992), as shown in Eq.(2):

$$N_S = \left(1-\left(1-S_c\right)^{1/\left(Se_s \times S_p \times N_A\right)}\right) \times \left(N_A - \left(S_p \times N_A - 1\right)/2\right) \quad (2)$$

where S_c is the confidence (95%) and S_p is the expected prevalence (5%).

Experimental data: Data from four experiments conducted at CIDC-Lelystad (Eblé et al., 2004; Eblé et al., 2007; Orsel et al., 2007) were used to estimate Se_c and Se_s, as well as P_{SL}. In these experiments 66 pigs were vaccinated of which 50 became infected[1] either by inoculation or via direct contact with infected pigs. Ten of these 50 pigs died or were euthanized in the first two weeks after infection and were therefore excluded from the analysis, leaving 40 pigs to estimate Se_c, Se_s, and P_{SL}.

All pigs were scored according to a list of clinical signs, both FMD-specific and non-specific. Only those clinical signs that indicated lameness (six parameters) and vesicles (five parameters) were used in the subsequent analysis. Twenty six of all 40 pigs had ≥ 2 clinical signs during ≥ 3 consecutive days and were assumed to have been detected on farm. These numbers were used to define an uncertainty distribution for Se_c (Vose, 2000). Of the 14 pigs not detected on farm, 3 had clinical signs associated with lameness (ante-mortem inspection) and/or vesicles (post-mortem inspection) and were assumed to have been detected at slaughter. These numbers were used to define an uncertainty distribution for P_{SL} (Vose, 2000). Furthermore, 33 of all 40 pigs were positive in the NSP-ELISA. These numbers were used to define an uncertainty distribution for Se_s (Vose, 2000).

Prevalence calculations

Bayesian inference (see, e.g. Vose, 2000) was used to estimate uncertainty distributions for P_H and P_A, assuming thatdespite at least one infected animal being present,all test results were negative.

Within-herd prevalence: The prior distribution of P_A was derived from final size calculations (de Jong & Kimman, 1994; Velthuis et al., 2007) predicted for an FMD outbreak in a vaccinated pig herd, assuming a basic reproduction ratio R_0 of 2.42 (Orsel et al., 2007), herd sizes as given by the model, and one infected pig at the start of the infection chain.

For each test protocol a likelihood distribution was calculated for the probability that only negative test results were obtained given ≥ 1 infected animals present in the herd. The likelihood distribution for *clinical* was calculated as in Eq.(3):

$$P\left(T^- \mid A_i = i\right) = \left(1 - Se_c\right)^i \quad (3)$$

where T^- = all test results negative, A_i = number of infected animals, and i = 1, 2, …, N_A. The likelihood distribution for *serology* was given by Eq.(4):

[1] A pig was considered infected if virus was isolated from oropharyngeal fluid.

$$P\left(T^- \mid A_i = i\right) = \left(1 - \frac{i}{N_A} \times Se_s\right)^{N_S} \quad (4)$$

and the the likelihood distribution for *waiting* was calculated as in Eq.(5):

$$P\left(T^- \mid A_i = i\right) = \left[\left(1 - Se_w\right)^i\right]^m \quad (5)$$

The posterior distribution of PA after final screening was calculated by multiplying the prior distribution by the likelihood distributions for clinical and serology. To obtain the posterior distribution of PA at the end of the six-month waiting period, the prior distribution was multiplied with all three likelihood distributions. In the model calculations PA was sampled from these posterior distributions.

Herd prevalence: Calculation steps for P_H were similar to those for P_A. In these calculations, however, an uninformed prior was used giving all possible values for P_H equal probability. Although this does not resemble reality – only a few infected herds will be missed in the final screening –, no quantitative estimate for the between-herd R_0 in a vaccinated pig population was available. What-if analysis indicated that using an $R_0 < 1$ for between-herd spread did not affect model results. Because each of the three testing protocols included all herds, their likelihood distributions were all based on Eq.(3) and calculated as:

$$P\left(T^- \mid H_i = i\right) = \left(1 - HSe_j\right)^i \quad (6)$$

where T– = all test results negative, Hi = number of infected herds, i = 1, 2, …, NH, and HSej = herd sensitivity of testing protocol j with j = clinical, serology, or waiting.

Herd sensitivity: Herd sensitivity (HSe) depends on the number of infected animals on the farm, the number of animals tested, and test sensitivity (Se) and specificity (Sp) (Martin et al., 1992). The HSe for each testing protocol was estimated using both the prior and likelihood distributions of PA. The prior distribution gives the probability that 1, 2, …, NA infected animals are present on the farm, whereas the likelihood distributions gives the probability that 1, 2, …, NA infected animals are not detected taking into account the number of animals tested and test Se and Sp, the latter being 1. The probability that a farm will not be detected is therefore given by Eq.(7):

$$P\left(T^- \mid A^+\right) = \sum_{i=1}^{N_A} p_i \times P\left(T^- \mid A_i\right) \quad (7)$$

where T^- means all test results negative, A^+ means ≥ 1 infected animals present, p_i is the probability that i infected animals are present based on the prior distribution of P_A, and $P(T^- \mid A_i)$ is the probability that all test results are negative given i infected animals on the farm (Eqs.(3)-(5)). The probability that a farm will be detected is thus given by Eq.(8):

$$P\left(T^+ \mid A^+\right) = 1 - Eq.7 \quad (8)$$

where T^+ means ≥ 1 positive test result, and equals *HSe*.

The prior distribution for P_A was dependent on R_0 and N_A, and thus differed for each herd size in the model. Furthermore, test sensitivities were modelled as uncertainty distributions, not fixed values. Therefore, calculations for *HSe* were performed for each herd size included in the model (16 values) using @Risk (10,000 iterations). The final values used for HSe_c, HSe_s, *and* HSe_w (see Table 1) were based on a weighted average of the mean values per herd size, taking into account the number of herds in each herd size class.

RESULTS

Within-herd prevalence

P_A is on average 8.3×10^{-3} after final screening and 7.3×10^{-3} at the end of the six-month waiting period. The prior distribution for P_A is bimodal, i.e., both minor and major outbreaks can occn the farm when $R_0 = 2.42$. Major outbreaks will almost certainly be detected, as indicated by the likelihood functions for *clinical*, *serology*, and *waiting*. In Figure. 2 the contributions of the prior and likelihood distributions to the posterior distributions for P_A are shown for the average herd size in the vaccinated areafor low values of P_A. The likelihood function for *serology* does not add much information (rather flat line), because only a sample of pigs is tested. The likelihood function for *waiting* has most influence and almost completely determines the shape of the posterior distribution for P_A at the end of the six-month waiting period.

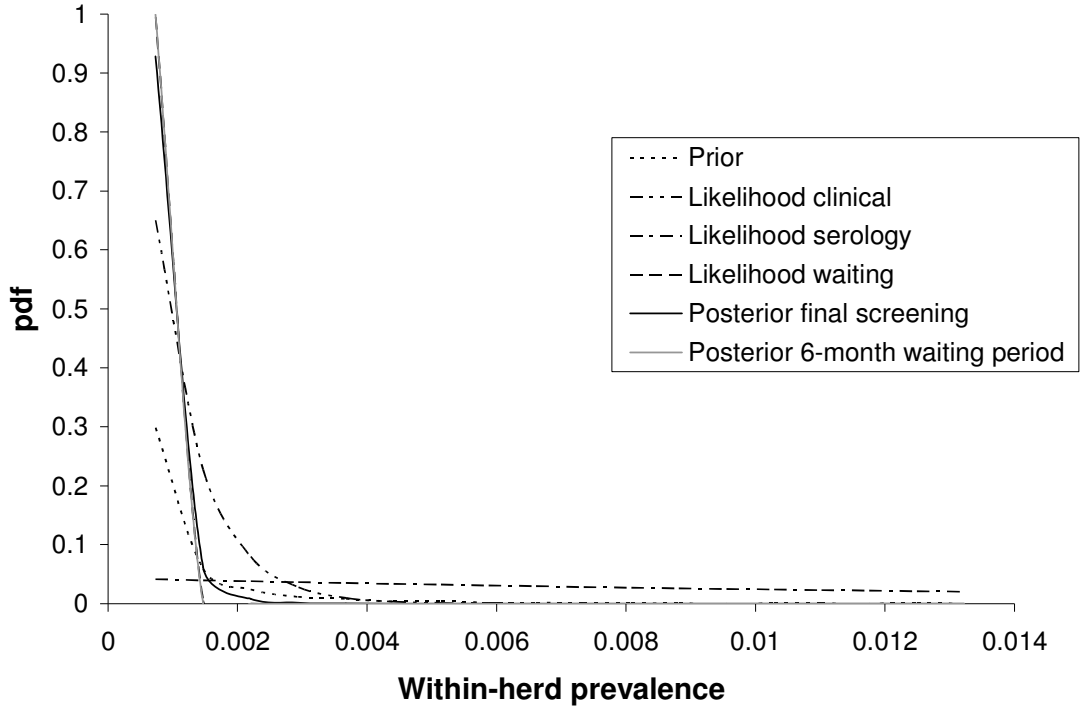


Fig. 2 Probability density function (pdf) of prior, likelihood, and posterior distributions of the within-herd prevalence for the average herd size in the vaccinated area (1361 pigs)

Herd prevalence

P_H is on average 3.3×10^{-3} after final screening and 3.1×10^{-3} at the end of the six-month waiting period. In Figure 3 the contributions of the prior and likelihood distributions to the posterior distributions for P_H are shownfor low values of P_H. The impact of a non-informative uniform prior distribution for P_H is very small. Now, the likelihood function for *serology* has more influence, because all herds are included in the serological testing. Again, the likelihood function for *waiting* has most influence and almost completely determines the shape of the posterior distribution for P_H at the end of the six-month waiting period.

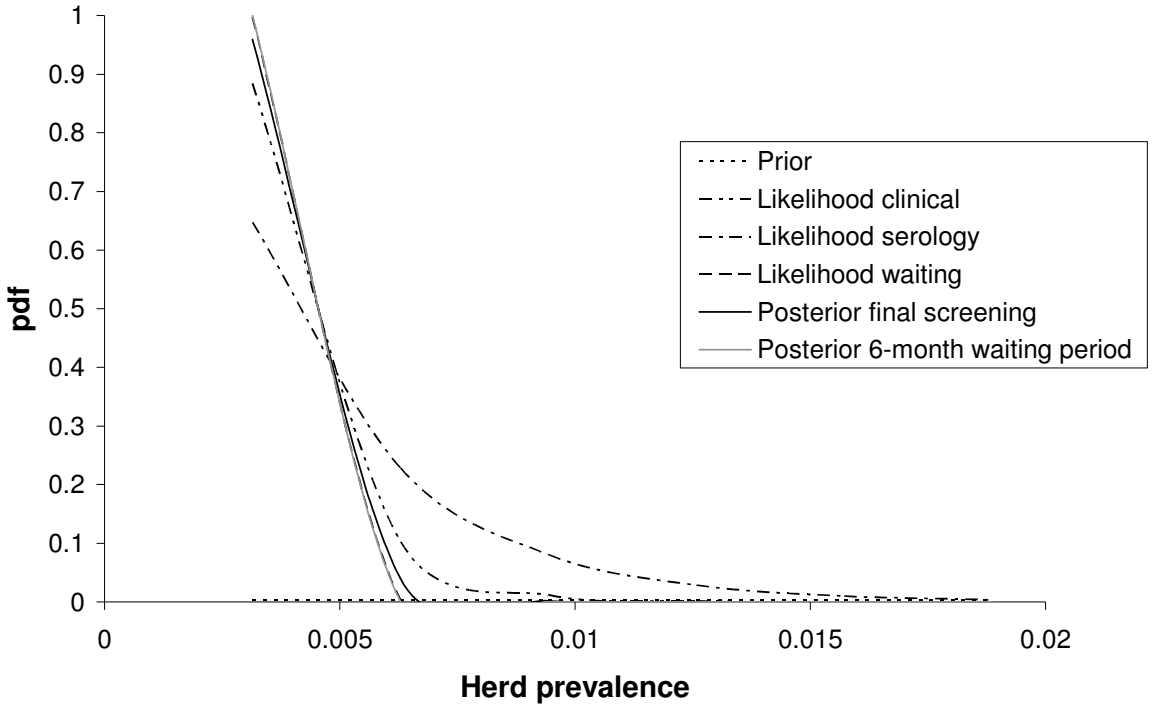


Fig. 3 Prior, likelihood, and posterior distributions of the herd prevalence in the vaccinated area (318 herds)

Probability of processing a carcass derived from an FMD-infected pig

Model calculations indicate that P_{carc} is on average 2.0×10^{-5} after final screening (Table 2). This probability is reduced to 1.7×10^{-5} after an additional waiting period of five months. Median values of P_{carc} are ~6 times lower, indicating that the distribution is skewed to the right. Figure 4 gives the cumulative density function of P_{carc} after final screening and at the end of the six-month waiting period. Differences are very small, except for the right tail of the distribution. The maximum calculated value of P_{carc} is 1.0×10^{-3} at the end of the six-month waiting period, whereas 7 out of 10,000 iterations return a higher probability for P_{carc} after final screening with a maximum value of 2.7×10^{-3}.

For comparison purposes model calculations were also run for a three-month waiting period (m = 2). As expected, P_{carc} at the end of a three-month waiting period is slightly lower than after the final screening and slightly higher than at the end of the six-month waiting period (Table 2).

Table 2. Model output for the probability that a carcass derived from an FMD-infected pig would be processed (P_{carc}) after final screening and at the end of a three- and six-month waiting period

	Mean	SD	0.05 Perc.	Median	0.95 Perc.	Maximum
Final screening	2.00×10^{-5}	1.10×10^{-4}	5.30×10^{-7}	2.97×10^{-6}	2.83×10^{-5}	2.72×10^{-3}
3-month waiting period	1.80×10^{-5}	9.86×10^{-5}	5.21×10^{-7}	2.87×10^{-6}	2.47×10^{-5}	1.80×10^{-3}
6-month waiting period	1.71×10^{-5}	9.32×10^{-5}	5.16×10^{-7}	2.82×10^{-6}	2.43×10^{-5}	1.01×10^{-3}

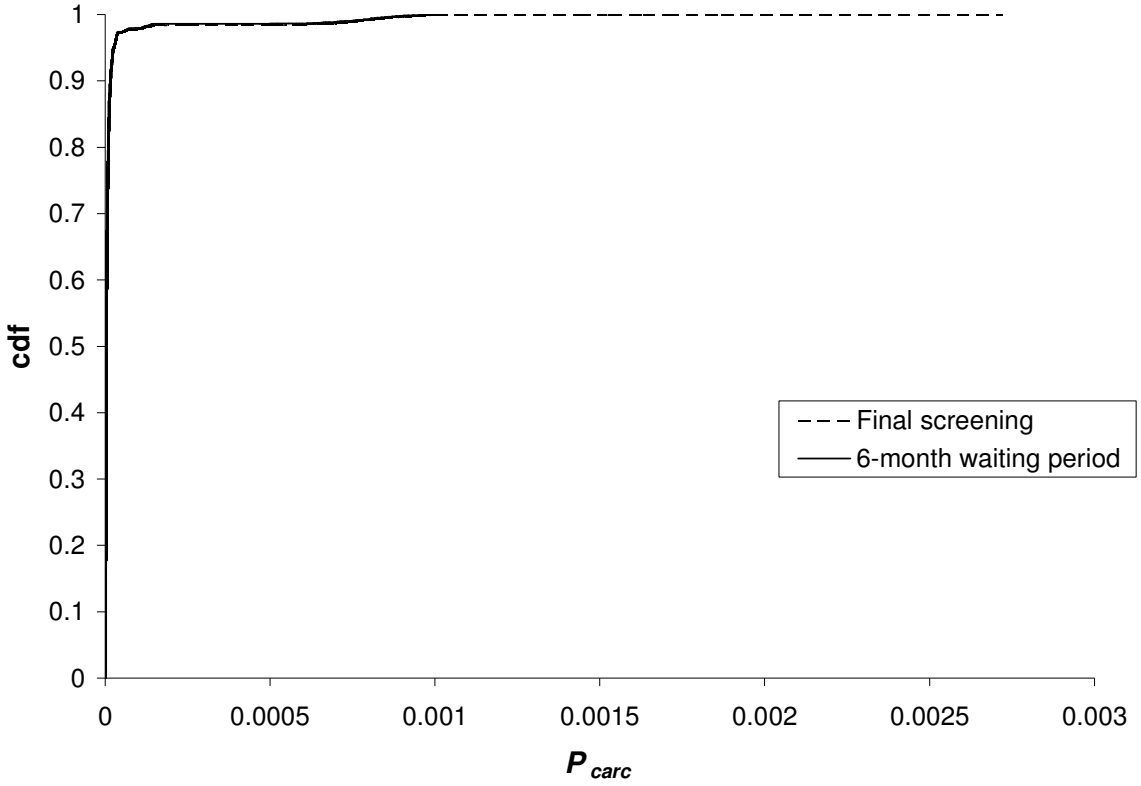


Fig. 4 Cumulative density function (10,000 iterations) of the probability that a carcass of an FMD-infected pig will be processed (P_{carc}) after final screening and at the end of a six-month waiting period

DISCUSSION

The calculated probability that a carcass of an FMD-infected pig will be processed is very low with an average value of 2.0×10^{-5} directly after final screening and 1.7×10^{-5} after a six-month waiting period. A six-month waiting period thus does not greatly reduce the risk of exporting FMD-infected carcasses. The calculated probabilities are in the same order of magnitude as the probabilities calculated for the FMD risk of importing deboned beef from vaccinated areas in South-America (Astudillo et al., 1997; Anonymous, 2002).

Despite these low probabilities, it cannot be excluded that some infected carcasses will be exported since the Netherlands is a major exporter of pork with an annual export of ~6.8×10^{5}

tons (CBS, 2007). Given a slaughtered weight of 90 kg per pig (KWIN, 2006), this equals ~7.6 million slaughtered pigs. Given that the vaccinated region harbours ~3.8% of all pigs in the Netherlands, 242 carcasses derived from FMD-infected pigs are expected to be exported during the waiting period after final screening (i.e. during a five-month period).

The calculated value of P_{carc} overestimates the risk associated with resuming export of pig carcasses after the final screening for three reasons. Firstly, model calculations only took into account screening results from pig farms. Cattle, sheep, and goat farms will, however, also be clinically inspected and serologically tested at the end of an epidemic, since FMD is a contagious disease affecting all cloven-hoofed animals. Only when all test results on all species are negative will the vaccinated area be declared free from FMD. Inclusion of all farms in the final screening will thus increase the confidence that the area is free as well as the probability that P_H is low (i.e. few farms infected) if the area is not truly free. Secondly, the model calculated the probability that a carcass is derived from an FMD-infected pig regardless of its infection status, i.e., P_H and P_A were based on both viraemic and seropositive pigs. Viraemia in pigs will last ~4-10 days (Farez & Morley, 1997), whereas serological response will last much longer. Only viraemic pigs and their carcasses pose an infection risk. Thirdly, carcasses derived from viraemic pigs can only result in new infections when ingested by other pigs. Most carcasses are, however, destined for human consumption. Furthermore, swill feeding is prohibited in the EU (CEC, 2001), further reducing the probability that pork will end up with pigs. Only illegal feeding of unheated swill can thus initiate new FMD infections.

The risk of exporting pig carcasses directly after final screening can be reduced further by appropriate risk management. FMDV is pH and temperature labile and will rapidly be inactivated when pH < 6 (Anonymous, 2006). Thermal inactivation is obtained with an internal temperature of ≥ 69°C (McKercher et al, 1980). Exporting only heated pork products thus reduces the risk. Deboning and maturation might, however, not suffice to inactivate FMDV in pork, because this does not naturally result in a pH < 6 as is seen with maturation of beef (Farez & Morley, 1997). Furthermore, not all tissues will contain equal amounts of FMDV. In pigs, the greatest quantities of virus are found in the blood, epithelium, and liver (Sellers, 1971). In particular, the skin contains high titres of FMDV (Alexandersen et al., 2001). Withholding bacon from export will thus further reduce the risk of exporting pork derived from infected vaccinated pigs.

The method used in this study did not answer the question of whether or not the affected area was correctly declared free from FMD, but only provided insight into the probability of processing carcasses from FMD-infected pigs if it was not. Indeed, absence of disease is impossible to prove except by continuous monitoring of all animals with a perfect test (de Koeijer, 2006). Most methods used to estimate the probability that an area is free from disease are based on the so-called design prevalence (see e.g. Cannon, 2002; Martin et al., 2007), giving the probability that disease is present at a level below this design prevalence (which includes the probability that it is not present at all). The outcome of this study is more informative because it gives full uncertainty distributions of both P_H and P_A when disease would still be present. The calculated values indicate that P_H is in all cases lower than the commonly used design P_H of 2% (see Table 2) (Greiner & Dekker, 2005). In other words,the affected region would have been declared free from disease with 100% certainty using methods based on design prevalence. Mintiens et al. (2007) developed a method to calculate the probability of true disease freedom evading the use of design prevalence, i.e., they calculated the probability that no infected animals are in fact present. This approach can only be used if test specificity < 1 and is based on a threshold number for positive test results, i.e., a region is declared free from disease if the

number of positive test results is lower than this threshold, assuming that these are all false positives. It is, however, doubtful if any trading partner will accept positive test results in declaring an area free from disease.

Although model calculations could not indicate the probability that the area is truly free from disease, they gave a good estimate of the probability that carcasses from FMD-infected pigs would be processed when a vaccinated area was incorrectly declared free after final screening. Resuming export after a six-month waiting period did not greatly reduce this probability.

REFERENCES

Alexandersen, S., Oleksiewicz, M.B. and Donaldson, A.I. (2001). The early pathogenesis of foot-and-mouth disease in pigs infected by contact: a quantitative time-course study using TaqMan RT-PCR. J. Gen. Virol. 82, 747-755

Anonymous (2002). Risk assessment – Importation of fresh (chilled or frozen) beef from Uruguay. USDA Animal and Plant Health Inspection Service. 48p

Anonymous (2006). Risk assessment on foot and mouth disease: Part I. Assessing the risk of foot and mouth disease introduction into the EU from developing countries. EFSA-Q-2004-113. The EFSA Journal 313

Astudillo, V., Sutmoller, P., Saraiva, V. and López, A. (1997). Risks of introducting foot and mouth disease through the importation of beef from South America. Rev. Sci. Tech. Off. Int. Epiz. 16, 33-44

Bouma, A., Elbers, A.R.W., Dekker, A., de Koeijer, A., Bartels, C., Vellema, P., van der Wal, P., van Rooij, E.M.A., Pluimers, F.H. and de Jong, M.C.M. (2003). The foot-and-mouth disease epidemic in The Netherlands in 2001. Prev. Vet. Med. 57, 155-166

Brocchi, E., Bergmann, I.E., Dekker, A., Paton, D.J., Sammin, D.J., Greiner, M., Grazioli, S., De Simone, F., Yadin, H., Haas, B., Bulut, N., Malirat, V., Neitzert, E., Goris, N., Parida, S., Sørensen, K. and De Clercq, K. (2006). Comparative evaluation of six ELISAs for the detection of antibodies to the non-structural proteins of foot-and-mouth disease virus. Vaccine 24, 6966-6979

Cannon, R.M. (2002). Demonstrating disease freedom – combining confidence levels. Prev. Vet. Med. 52, 227-249

CBS (2005). Landbouwtelling 2005 (in Dutch). http://statline.cbs.nl/StatWeb/start.asp?LA=nl&DM=SLNL&lp=Search%2FSearch

CBS (2007). Statline databank (in Dutch). http://statline.cbs.nl/StatWeb/start.asp?LA=nl&DM=SLNL&lp=Search%2FSearch

CEC (Commission of the European Communities) (2001). Council Directive 2001/89/EC of 23 October 2001 on Community measures for the control of classical swine fever (Text with EEA relevance). Off. J. Eur. Communities L 316, 5-35

CEC (Commission of the European Communities) (2003). Council Directive 2003/85/EC of 29 September 2003 on Community measures for the control of foot-and-mouth disease repealing Directive 85/511/EEC and Decisions 89/531/EEC and 91/665/EEC and amending Directive 92/46/EEC (Text with EEA relevance). Off. J. Eur. Union L 306, 1-87

Eblé, P.L., Bouma, A., de Bruin, M.G.M., van Hemert-Kluitenberg, F., van Oirschot, J.T. and Dekker, A. (2004). Vaccination of pigs two weeks before infection significantly reduces transmission of foot-and-mouth disease virus. Vaccine 22, 1372-1378

Eblé, P.L., Bouma, A., Weerdmeester, K., Stegeman, J.A. and Dekker, A. (2007). Serological and mucosal immune responses after vaccination and infection with FMDV in pigs. Vaccine 25, 1043-1054

Farez, S. and Morley, R.S. (1997). Potential animal health hazards of pork and pork products. Rev. Sci. Tech. Off. Int. Epiz. 16, 65-78

Greiner, M. and Dekker, A. (2005). On the surveillance for animal diseases in small herds. Prev. Vet. Med. 70, 223-234

de Jong, M.C.M. and Kimman, T.G. (1994). Experimental quantification of vaccine-induced reduction in virus transmission. Vaccine 12, 761-766

de Koeijer, A.A. (2006). Freedom of disease: a theoretical problem with practical issues. In: Dekker, A. and Nodelijk, G. (Eds.) Proceedings of the Dutch Society for Veterinary Epidemiology and Economics, Lelystad, 16 February 2006, pp. 7-8

KWIN (2006). Kwantitatieve Informatie Veehouderij 2006-2007 (in Dutch). Animal Sciences Group, Wageningen UR, Lelystad. 418p

Martin, S.W., Shoukri, M. and Thorburn, M.A. (1992). Evaluating the health status of herds based on tests applied to individuals. Prev. Vet. Med. 14, 33-43

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

McKercher, P.D., Morgan, D.O., McViar, J.W. and Shuol, N.J. (1980). Thermal processing to inactivate viruses in meat products. Proc. Ann. Meet. U.S. Anim. Hlth Assoc. 84, 320-328

Mintiens, K., Balavarca, Y., Verloo, D. and Aerts, M. (2007). Sample sizes for declaring disease freedom with uncertain diagnostic test performance and prevalence. In: Mellor, D.J. and Newton, J.R. (Eds.) Proceedings of the Society for Veterinary Epidemiology and Preventive Medicine, Dipoli, Helsinki/Espoo, Finland, 28-30 March 2007, pp. 29-37

OIE (2007). Terrestrial Animal Health Code. http://www.oie.int/eng/normes/mcode/A_summry.htm?e1d11

Orsel, K., de Jong, M.C.M., Bouma, A., Stegeman, J.A. and Dekker, A. (2007). Foot and mouth disease virus transmission among vaccinated pigs after exposure to virus shedding pigs. Vaccine 25, 6381-6391

Palisade Corporation (2002). Guide to using @Risk. Risk Analysis and Simulation Add-In for Microsoft Excel, Version 4.5. Palisade Corporation, Newfield, NY. 499p

Sellers, R.F. (1971). Quantitative aspects of the spread of foot and mouth disease. Vet. Bull. 41, 431-439

Tomassen, F.H.M., de Koeijer, A.A., Mourits, M.C.M., Dekker, A., Bouma, A. and Huirne, R.B.M. (2002) A decision-tree to optimise control measures during the early stage of a foot-and-mouth disease epidemic. Prev. Vet. Med. 54, 301-324

Velthuis, A.G.J., Bouma, A., Katsma, W.E.A., Nodelijk, G. and de Jong, M.C.M. (2007). Design and analysis of small-scale transmission experiments with animals. Epidemiol. Infect. 135, 202-217

Vose, D. (2000). Risk analysis. A quantitative guide, Second edition. John Wiley & Sons, Chichester. 418

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, Eire	Collins
1999	Warwick	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

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

EXECUTIVE COMMITTEE 2007-2008

J.R. Newton (President), E.J. Peeler (Senior Vice-President), L. Alban (Junior Vice-President), T. Parkin (Honorary Secretary), L. Kelly (Honorary Treasurer), S. More, G.L. Pinchbeck, D. U. Pfeiffer, L. Alban, T.D.H. Parkin, S. Robinson (Co-opted)

Honorary Auditors: Dominic Mellor & Fraser Menzies

LIFE MEMBERS

J.M. Booth, M.J. Clarkson, J.D Collins, G. Davies, J.T. Done R.G. Eddy, P.R. Ellis, E.A. Goodall, M.E. Hugh-Jones, W. Martin, F. Menzies, A.M. Russell, M.V. Thrusfield

PLENARY TALKS

Year	Gareth Davies Lecture	Conference Opening Plenary
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 (£20 sterling) 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 web site, http://www.svepm.org.uk/, or from the Secretary or Treasurer.

Please send this form to the Society's Treasurer:

Dr Louise Kelly
Department of Statistics and Modelling Science
University of Strathclyde
Glasgow
G1 1XH

☎ **+44 (0) 141 548 3659**
FAX +44 (0) 141 552 2079
Email: louise@stams.strath.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