Available online at www.sciencedirect.com
ScienceDirect
The Veterinary Journal xxx (2007) xxx–xxx
The Veterinary Journal
www.elsevier.com/locate/tvjl

An investigation of factors associated with the prevalence of verocytotoxin producing *Escherichia coli* O157 shedding in Scottish beef cattle

G.J. Gunn [a,*], I.J. McKendrick [b], H.E. Ternent [c,1], F. Thomson-Carter [d,2], G. Foster [b], B.A. Synge [b,3]

[a] *SAC Animal Health Group, Epidemiology Research Unit, Stratherrick Road, Inverness IV2 4JZ, UK*
[b] *Biomathematics and Statistics Scotland, Kings Buildings, Edinburgh EH9 3JZ, UK*
[c] *SAC Veterinary Services, Drummondhill, Stratherrick Road, Inverness IV2 4JZ, UK*
[d] *Scottish E. coli Reference Laboratory, Grampian University Hospitals Trust, Aberdeen AB25 2ZN, UK*

Accepted 21 August 2007

Abstract

The prevalence of verocytotoxin-producing *Escherichia coli* (VTEC) O157 in 12–30-month-old beef finishing cattle in Scotland was determined using 1 g faeces samples enriched in buffered peptone water, followed by immunomagnetic separation (IMS) and isolation on sorbitol MacConkey agar with cefixime and tellurite supplement (CT-SMAC). A validated questionnaire was used to collect information that could be associated with the samples. Generalised Linear Models and Generalised Linear Mixed Models were used to identify factors associated with shedding both between and within groups.

A total of 14,856 samples were collected from 952 farms, of which 1231 were positive for VTEC O157. Prevalence levels were calculated with 95% confidence intervals as follows: 7.9% (6.5%, 9.6%) of animals sampled were estimated to be shedding VTEC O157, while 22.8% (19.6%, 26.3%) of farms were estimated as having at least one animal shedding in the group sampled. The median percentage of animals shedding in positive groups was 25% (20%, 32%).

An increased probability of a group containing a shedding animal was associated with larger numbers of finishing cattle, the presence of pigs on the farm, or the farm being classed as a dairy unit stocking beef animals. Farms that spread slurry on grazing land were more likely to have shedding animals, while those that spread manure were at lower risk. Groups with older animals were less likely to be identified as positive. There was no significant regional difference in group shedding probabilities, but the proportion of positive groups dropped over two successive years of the study. Higher mean levels of shedding in positive groups were associated with animals being housed rather than at pasture, and this effect was stronger in groups which had recently had a change in housing or diet. Farms with animals at pasture had lower mean prevalence where water was supplied from a natural source, as had farms with higher numbers of finishing cattle. There remained unexplained variability in mean prevalence levels on positive farms in different areas of Scotland.

Keywords: *E. coli*; VTEC O157; Cattle; Shedding prevalence; Risk factors

Introduction

Verocytotoxin-producing *Escherichia coli* O157 is an important cause of diarrhoea in man and in some cases serious consequences follow, such as haemorrhagic colitis, haemolytic uraemic syndrome or thrombocytopaenia,

* Corresponding author. Tel.: +44 1463 243030; fax: +44 1463 711103.
E-mail address: Goerge.Gunn@sac.ac.uk (G.J. Gunn).
[1] Present address: Veterinary Comparative Epidemiology and Informatics, University of Glasgow Veterinary School, Glasgow G61 1QH, UK.
[2] Present address: Institute for Environmental Science and Research, P.O. Box 50-384, Porirua, Wellington, New Zealand.
[3] Present address: SAC Veterinary Services, Allan Watt Building, Bush Estate, Penicuik, Midlothian EH26 0QE, UK.

1090-0233/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tvjl.2007.08.024

which may be fatal (Karmali, 1989). There are approximately 200 human cases reported annually in Scotland and the rate per unit population is approximately three times that in England and Wales (Smith et al., 1998).

Although human infection may arise from person to person contact and from consumption of food contaminated by asymptomatic human carriers, it is accepted that several species of animal (cattle, sheep, horses, goats, dogs and geese) also carry the organism (Møller Nielsen et al., 2004). Frequently, primary human infection can be attributed to contamination of the environment or the food chain from such animals, in particular cattle (Riley et al., 1983; Espie et al., 2006), but a robust estimate for the prevalence of shedding of the organism from Scottish cattle has not been established. There remains little understanding of the factors that might influence shedding. A greater understanding of the epidemiology of VTEC O157 could lead to possible interventions at farm level to reduce the shedding of this hazardous organism.

Published information about the prevalence of faecal shedding of *E. coli* O157 in cattle populations shows wide variation. Table 1 lists examples of reported animal level prevalence from different classes of cattle, using different tests, in different countries. True comparisons are difficult because the execution of these studies has not been consistently rigorous and problems arise from the lack of uniformity of the study designs and the laboratory methods. Outcomes will have been affected by the populations from which samples are drawn; the sampling methodologies adopted; the timing of sampling; and the sensitivity of the tests used for screening. An early prevalence study in Scotland (Synge and Hopkins, 1996) using direct plating on SMAC agar found only 0.25% of bovine faeces samples submitted to veterinary investigation laboratories positive for VTEC O157. Another study of cattle at the Sheffield abattoir, using similar methods, found 4% of cattle to be shedding the organism (Chapman et al., 1993). The introduction of immunomagnetic separation (IMS) as a more sensitive technique (Chapman et al., 1994) provides a tool for the generation of more accurate estimates of prevalence, reducing the downwards bias of estimates from direct plating methods.

This study was designed with the objective of estimating the herd level prevalence VTEC O157 shedding for fattening cattle in Scotland using IMS test on 1 g faeces samples. As sub-objectives, the animal level prevalence was to be estimated and the effect of a variety of potential risk factors on these prevalences investigated.

Materials and methods

Study design

Herds likely to contain fattening cattle were selected randomly using a sampling frame derived from the Scottish Executive farm census. For the purpose of the study Scotland was divided into six regions: the five Animal Health Divisions plus the Northern and Western Isles forming a separate region (Fig. 1).

The set of farms to be sampled was stratified by farm-management type and by region. Pilot data derived using the proposed sampling protocol, suggesting a mean within-farm animal prevalence of 10%, was used to derive an on-farm sampling plan aiming to identify as positive 80% of groups containing at least one shedding animal. Assuming a herd prevalence of 2%, the number of herds to be sampled was specified as giving an 80% probability of the 95% confidence interval for the farm prevalence derived from the sampling protocol being bounded by a tolerance of ±1% around the true value.

On the basis of these assumptions, it was calculated that the study was unlikely to have an acceptable statistical power to detect anything other than major differences in prevalence caused by potential risk factors. Nevertheless, it was thought worthwhile to collect information on potential risk factors from the farms included in the survey. However, since the herd prevalence of shedding proved to be substantially higher in practice, it became possible to carry out an extensive investigation into risk factors using information collected from the farm management questionnaire.

Field procedures

Each region was sampled in turn on a weekly rotation, with sample herds being visited on a random basis within each region. Only cattle aged 12–30 months were sampled and on each farm the group of such animals nearest to slaughter was selected. At each visit a farm management questionnaire was completed and faecal samples were collected and returned to the laboratory for analysis.

Faecal pat sampling

When collecting samples from faecal pats, the pats typically cannot be identified as coming from any specific animal. Assuming that shedding animals do not defaecate at any increased rate relative to non-shedding animals, sampling without replacement from the population of faecal pats is equivalent to sampling with replacement from the population of animals, provided that there is a moderately sized population of faecal pats. Pilot data suggested a within-herd prevalence that was distributed as a beta distribution with a mean of approximately 10%. A

Table 1
Estimates of *E. coli* O157 animal level prevalence from the published literature

Country	Class of cattle	Isolation method	Prevalence (%)	Reference	Year
Scotland	Diagnostic samples	Direct culture	0.25	Synge and Hopkins (1996)	1993
Sheffield, UK	Abattoir	Direct culture	4	Chapman et al. (1993)	1993
USA	Dairy cattle	Enrichment	0.28	Hancock et al. (1994)	1994
USA	Feedlot cattle	Direct culture	0.33	Hancock et al. (1994)	1994
USA	Calves	Enrichment	1.5	Zhao et al. (1995)	1995
England and Wales	Diagnostic samples	IMS	0.83	Richards et al. (1998)	1998
Netherlands	Adult cattle	IMS	11.1	Heuvelink et al. (1996)	1996
Finland	Slaughter cattle	IMS	1.31	Lahti et al. (2001)	2001

Please cite this article in press as: Gunn, G.J. et al., An investigation of factors associated with the prevalence of ..., The Veterinary Journal (2007), doi:10.1016/j.tvjl.2007.08.024

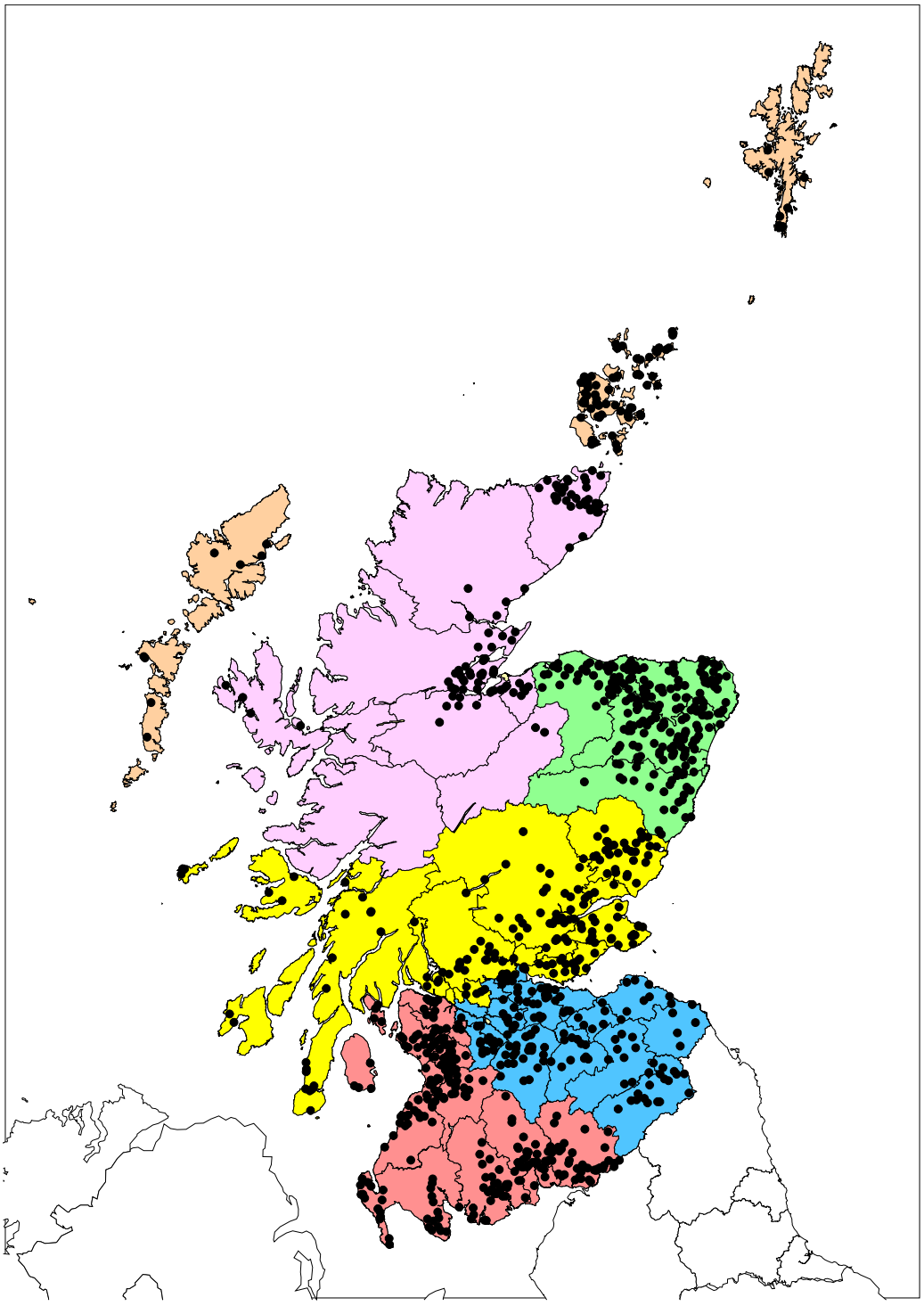

Fig. 1. The six regions of Scotland used for the study. (brown = Islands, purple = Highlands, green = North East, yellow = Central, blue = South East and pink = South West. The dots represent the positions of the herds sampled. (This work is based on data provided with the support of the ESRC and JISC and uses boundary material which is copyright of the Crown, and the Post Office. *Source:* The 1991 Census, Crown Copyright. ESRC purchase). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

beta-binomial distribution was used to model the number of positive animals in the herd.

From this model the numbers of samples required to provide an 80% probability of detecting that a herd contains shedding cattle were calculated. Samples were collected from freshly voided faecal pats into sterile plastic universal containers, returned to the laboratory and tested within 48 h.

Farm management questionnaire

At each visit, in addition to taking faecal samples, a farm management questionnaire was completed. There were two operators who collected samples over the course of the study. A detailed validation procedure was undertaken to validate the questionnaire and eliminate operator bias. Farm personnel were asked questions on the numbers of cattle on the farm

Please cite this article in press as: Gunn, G.J. et al., An investigation of factors associated with the prevalence of ..., The Veterinary Journal (2007), doi:10.1016/j.tvjl.2007.08.024

and about the numbers of groups kept, their source and breed type. For housed cattle questions were asked regarding the type of housing and the management during the previous 4 weeks. This included the timing of housing or movement, the type of forage or concentrate fed and how the silage was made. For grazing animals questions were asked about movements of the group, applications of slurry or manure onto the fields and supplementary feeding. For all groups of animals questions were asked about the presence of other animals on the farm and the water supply.

Laboratory procedures

(a) Isolation: One gram faeces samples were tested for *E. coli* O157 using standard methodology (Foster et al., 2003) comprising: enrichment in buffered peptone water (BPW) (Oxoid) without antibiotic supplement; immuno-magnetic separation with anti-*E. coli* O157 coated magnetic beads (Dyna-beads anti-*E. coli* O157, Dynal); resuspension and inoculation on to CT-SMAC plates incubated at 37 °C. After 18–24 h pale non-sorbitol fermenting colonies were tested for agglutination with *E. coli* O157 latex (Oxoid). A sub-culture from each latex positive sample was sent for confirmation and phage typing to the Scottish *E. coli* Reference Laboratory, Department of Medical Microbiology, Foresterhill, Aberdeen, Scotland. All staffs were validated in the techniques to rule out any inter-operator bias.

(b) Typing: All isolates were first confirmed as *E. coli* O157. They were phage typed using standard methods (Khakria et al., 1990) and examined for genes encoding the production of the Verocytotoxins VT1 and VT2 and the eae gene which encodes for enterocyte attachment and effacement using a multiplex PCR (Louie et al., 1994; Pollard et al., 1990).

Data management

A sample (10%) of records was checked against the original questionnaires. A 0.4% error rate was detected, with a 95% confidence interval of (0.2%, 0.6%), indicating that, on average, the data entry process had been reliable. These errors were, however, concentrated in a small number of data fields, and these were checked against the original questionnaire for all records. No entry errors were found in the recorded numbers of VT+ samples. All identified errors in the data were corrected prior to analysis.

Statistical analysis

If VTEC O157 was isolated from any pat from a group, the group was defined as positive. Confidence intervals for binomial proportions were calculated using the exact method (Armitage et al., 2002).

The prevalence data were initially analysed using a Generalised Linear Model (GLM), but the poor fit given by this approach necessitated the use of an alternative methodology. The data were treated as being the outcome of a mixture distribution, where a proportion p_{neg} of the population are defined as negative farms and will always return a zero number of positive samples. In the positive population, the between farm variability was modelled as a beta distribution, taking parameters a and b, while the sampling distribution of the faecal pat sampling process was taken to be binomial. Because of the strong negative correlation between p_{neg} and a and b, p_{neg} was set equal to the maximum likelihood estimate. Confidence intervals for prevalences were generated using the χ^2 approximation to the profile log-likelihood ratio. The distribution of number of cattle in the sampling groups was modelled as a log-normal distribution.

Risk factors were analysed using the GLM and Generalised Linear Mixed Model (GLMM) procedures in Genstat. The fitting of a single model to the dataset was not appropriate, since the large number of negative observations, coupled with a subset of observed farms with high shedding levels, gave rise to a badly fitting model. This reflects the bimodal nature of the dataset, where many farms are likely to be consistently negative for shedding. The subset of the data with non-zero shedding results was analysed, as was the entire dataset, restructured to record absence/presence of shedding at the farm level. The use of random effects in such models ensured that the highly unbalanced nature of the data did not give rise to biased estimates of epidemiological effects. However, the use of mixed models is highly computer intensive, and so GLMs were used for the univariate aspects of the analysis. GLMs and GLMMs were fitted with a Binomial distribution and a logit link function. Where appropriate, GLMs were fitted with an overdispersion parameter to model excess variability in the data. When fitting a GLMM, farm, veterinary practice and county were all examined as possible random effects. County and veterinary practice were not found to be useful in explaining any of the variability of the data, so farm was used as the sole random effect at the level of the datum. Other epidemiological variables were fitted as fixed effects.

GLM analyses were initially carried out in a univariate basis, analysing the marginal effect of each factor or variate in turn. When analysing the shedding levels in positive groups, it was concluded that much of the variability in the data was explained by a specific factor. A further set of restricted multi-factor analyses were therefore carried out, fitting a GLM model incorporating this factor in interaction with each candidate factor or variate in turn. The statistical significance of each term in the GLM was assessed using the change in model deviance.

Any factor or variate with a P-value $\leqslant 0.1$ in any of these analyses was carried forward for more detailed analysis. Groups of highly correlated candidate factors were assessed for goodness-of-fit, and those factors giving good fits were then reviewed using a forward stepwise selection algorithm and the Akaike information criterion to select candidates for inclusion in the multi-factor model. A forward stepwise selection algorithm was used to review whether any previously rejected factors or variates should be added to the draft multi-factor model. The resulting draft multi-factor model was then fitted using a GLMM. The statistical significance of each term in the GLMM was assessed using the χ^2 approximation to the Wald test. The study design factors animal health division and farm management type were included in the multivariate models, as were descriptive factors defining the time of sampling. All reported P-values are two-sided. Although multiple hypothesis tests have been applied to each factor during the univariate analyses and again during the development of the multi-factor models, the P-values have not been adjusted for repeated testing.

Results

Prevalence

Of the 14,856 samples collected from 952 herds, 1296 were positive for *E. coli* O157 and 1231 positive for VTEC O157. These VT-positive samples were sourced from 207 farms. Hence, the raw figures indicate that 21.7% of groups sampled contained shedding animals, and that the animal level prevalence is 8.3%. Using the beta-binomial model, it was estimated that the prevalence of VTEC O157 shedding in finishing cattle was 7.9% with a 95% confidence interval of (6.5%, 9.6%) and that 22.8% of finishing groups contained at least one positive shedding animal, with a 95% confidence interval of (19.6%, 26.3%).

The point-estimate and confidence interval for the group prevalence are both slightly higher than the raw estimates given earlier, since the model estimates implicitly adjust for farms with low shedding rates being misclassified as negative due to sampling variability.

Verocytotoxins

The vast majority (1168) of isolates expressed the VT2 gene only (94.9%, with a 95% confidence interval ranging

Please cite this article in press as: Gunn, G.J. et al., An investigation of factors associated with the prevalence of ..., The Veterinary Journal (2007), doi:10.1016/j.tvjl.2007.08.024

from 93.5% to 96.0%). Three isolates, or 0.2% of the total (0.05%, 0.7%) expressed VT1 only while 60 (4.9% with a 95% confidence interval ranging from 3.7% to 6.2%) expressed both VT1 and VT2.

Enterocyte attachment and effacement

Genes encoding enterocyte attachment and effacement (eae) were detected in all isolates including those that were non-toxigenic, giving a 95% confidence interval for the true prevalence (99.7%, 100%).

Analysis of the absence/presence of shedding

The majority of groups (78%) had no shedding animal detected. The results of the univariate analysis are summarised in Table 2. Values are given for all factors or variates which gave rise to a P-value $\leqslant 0.1$ in the univariate analysis, which were ultimately included in a multi-factor model, or which reflect aspects of the sampling design or of field or laboratory practice.

Several factors summarise information about the number of cattle on the farm; these are highly correlated, and ultimately only the numbers of animals in the sampling group, on the farm and in the finishing group were included in the draft multi-factor model. The biosecurity factor is derived from the animal sourcing factor, and is preferred in the later analysis. The breed factor summarises only whether or not a farm stocks animals from the Beef/Dairy/Dairy-Beef class. There is clear evidence of temporal changes in prevalence, with evidence of a year-on-year drop in prevalence. This result explains the apparent laboratory operator effect: the availability of different laboratory staff was aliased with time. Restricting attention only to time periods when members of staff were assaying comparable samples, there is no statistically significant evidence ($P = 0.77$) of any difference in the prevalences arising from their work. It is found that the breed factor lacks significance when other terms are included. This probably reflects the fact that the class of animal which the univariate model identified as associated with higher prevalence only occurred on six farms in total.

Number of cattle and biosecurity lacks significance when other cattle-related factors are included. The latter result can be explained by the observation that larger farms tend to buy in replacement cattle; only small operations will breed all replacements. Maximum age of animal is evaluated, since it shows evidence of statistical significance in some multi-factor models. The results of the multi-factor model are summarised in Table 3.

Table 2
Key results from the univariate analysis: absence/presence of shedding

Factor/Variable	P-value	Comments
Management type	0.80	'Beef' and 'Other' farms have higher mean probabilities than 'Dairy'
Animal Health Division	0.16	'Highland' farms have lower mean probabilities than others
Sampling month	**0.06**	Lower mean probabilities in January and February. Anomalously low mean probabilities in April and June, anomalously high mean probability in November
Sampling year	**0.004**	Consistent drop in mean probability with time
Sampling method	0.28	Lower mean probability for farms assessed using rectal samples
Field operator	0.18	Farms with samples collected by one operator had a higher mean probability than those with samples collected by another
Laboratory operator	**0.04**	Farms with samples assayed by one operator had a lower mean probability than farms assayed by two others
Number of finishing cattle	**<0.001**	Farms with between 50 and 199 finishing cattle have higher mean probabilities than those with <50 animals, the mean probability for farms with >200 animals is higher still
Number of management groups	**0.08**	More groups are associated with a higher mean probability
Number of animals in sampling group	**<0.001**	Farms with <11 animals in the sampling group have lower mean probabilities than those with 11–28 animals. Farms with >28 animals have still higher mean probabilities
Maximum age of animals in sampling group	0.31	Higher maximum age associated with lower mean probability
Source of animals	**0.01**	Farms classed as 'Buy in' and 'Both' show higher mean probabilities than those classed as 'Breeding only'
Biosecurity of farm	**0.03**	Farms classed as 'Open' show higher mean probabilities than those classed as 'Closed'.
Breed	**0.03**	Farms with stock classed as Beef/Dairy/Dairy-Beef have higher mean probabilities than others. No consistent pattern
Whether manure is spread on pasture	**0.02**	Farms with unhoused animals which also report the use of manure on grass have a lower mean probability than all other farms
Whether slurry is spread on pasture	**<0.001**	Farms with unhoused animals which also report the use of slurry on grass have a higher mean probability than all other farms
Number of cattle	**0.002**	Farms with <100 cattle have lower mean probabilities than those with more animals
Whether pigs are on farm	**0.01**	Farms with pigs have higher mean probabilities than those without
Whether farm is a dairy unit stocking beef animals	**0.02**	Farms of this type have a higher mean probability than other farms

Factors with P-values <0.1 in the univariate analysis are given in bold type.

Please cite this article in press as: Gunn, G.J. et al., An investigation of factors associated with the prevalence of ..., The Veterinary Journal (2007), doi:10.1016/j.tvjl.2007.08.024

Table 3
Results of the multi-factor analysis: absence/presence of shedding

Factor/variable	Effect	Log-odds ratio	SE	*P*-value
Sampling year	Allowing for the explanatory factors, farms sampled in year 1999 are at lower risk of being positive than those sampled in 1998	−0.425	0.21	0.04
	Allowing for the explanatory factors, farms sampled in year 2000 are at lower risk of being positive than those sampled in 1999	−0.371	0.26	0.15
	Allowing for the explanatory factors, farms sampled in year 2000 are at lower risk of being positive than those sampled in 1998	−0.795	0.31	0.01
Sampling month	A broad cyclical effect, with unexplained prevalence peaking in Summer and troughing in Winter. Anomalous changes in prevalences observed in a number of months, such as April, June and November	na	na	0.02
Number of animals in	Farms with 12–28 animals are at a higher risk of being positive than those with <12 animals	0.687	0.23	0.003
	Farms with >28 animals are at a higher risk of being positive than those with 12–28 animals	0.462	0.19	0.03
Number of finishing cattle	Farms with 50–199 animals are at a higher risk of being positive than those with 1–49 animals	0.367	0.19	0.05
	Farms with >200 animals are at a higher risk of being positive than those with 50–199 animals	0.614	0.30	0.04
Whether slurry is spread on pasture	Considering only farms with animals at pasture, those which spread slurry are at a higher risk of being positive than those which do not	1.205	0.32	<0.001
Whether manure is spread on pasture	Considering only farms with animals at pasture, those which spread manure are at a lower risk of being positive than those which do not	−1.155	0.36	0.001
Whether farm is a dairy unit stocking beef animals	Such farms are at a higher risk of being positive than other farms	1.965	0.64	0.002
Whether pigs are on farm	Farms with pigs are at a higher risk of being positive than those without pigs	0.892	0.35	0.01
Maximum age of animals in sampling group	Higher maximum age is associated with a lower risk of the farm being positive	−0.031	0.015	0.04

The number of animals in the sampling group is correlated with the number of samples collected from the group. A positive association could have been generated through the higher group sensitivity arising from a larger sample. Consideration of the data suggests that this is unlikely, but even if the result is discounted on this basis, the inclusion of the number of finishing cattle in the multi-factor model (even in the presence of the sampling group factor) indicates that the size of enterprise remains a highly significant risk factor.

No statistically significant geographical or management system variability was observed in either the univariate or multi-factor model. By contrast, the initial analysis showed evidence of a temporal trend towards lower prevalences, and this trend remained in the multi-factor model, unaffected by the explanatory factors. When included in the full multi-factor model, an effect of month is found to be significant. In particular, low mean prevalences are apparent in April and June, and a high mean prevalence in November. This pattern matches that observed in the univariate model: the effect is not an artefact of a poorly fitting model. Hence, it can be concluded that the farm level prevalence estimates do vary with month, in a fashion which is unaffected by the explanatory factors.

Analysis of levels of shedding in positive groups

It was notable that most herds that were identified as containing at least one shedding animal had a low proportion of positive pats. However, a spectrum of levels of shedding was seen and a number of herds appeared to have virtually all animals shedding (Fig. 2). The median percentage of shedding animals was 25% with a 95% confidence interval of (20%, 32%).

The results of the univariate analyses indicated that housing status had an overwhelming effect on mean prevalence. Many of the factors investigated in the study were partially or wholly aliased with housing. Hence, rather than reporting the simple univariate results, it is more informative to present the results from fitting each factor and an interaction term, in turn, to a model already containing housing status. These restricted multi-factor results are summarised in Table 4. Values are given for all factors or variates which gave rise to a *P*-value <0.1 in the univariate analysis, which were ultimately included in a multi-factor model, or which reflect aspects of the sampling design or of field or laboratory practice.

A highly statistically significant seasonal effect was observed in the marginal analysis of the sampling month factor, with mean prevalence levels peaking in January while being relatively low between June and October. This pattern matches the management of housing in Scottish herds, and when the housing factor was included in the model, the temporal factors exhibited no statistically significant effects. Highly statistically significant differences were identified between different animal health divisions. Several factors summarised information about the number of cattle on the farm; these were highly correlated, and ultimately

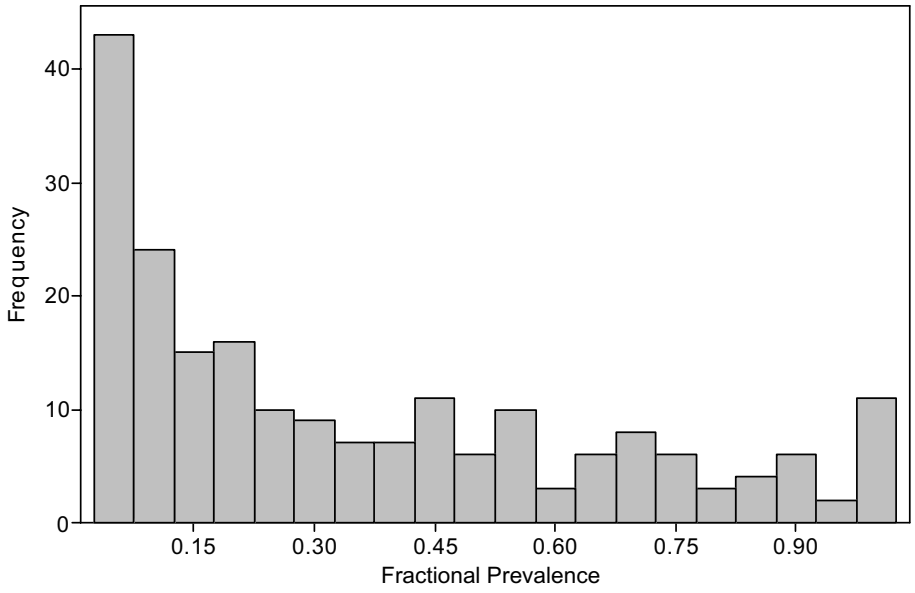

Fig. 2. Histogram of proportions of faecal pats positive for VTEC O157 from positive groups.

Table 4
Key results of the restricted multivariate analysis: levels of shedding in positive groups

Factor/variable	*P*-value	Comments
Management type	0.33	'Beef' farms have a higher mean prevalence and 'Others' a lower mean than 'Dairy' farms
Animal Health Division	**0.007**	'Highland' farms have a significantly higher mean prevalence than others
Sampling month	0.31	No apparent pattern in mean prevalence
Sampling year	0.23	No apparent pattern in mean prevalence.
Field operator	0.42	No apparent pattern in mean prevalence
Laboratory operator	0.45	No apparent pattern in mean prevalence
Number of finishing cattle	**0.032**	The larger the group of cattle, the lower the mean prevalence
Source of animals	**0.09**	Source significant in interaction with Housed factor. For unhoused animals, farms which 'Buy In' have lower mean prevalences, while for housed animals, farms which 'Buy In' have a higher mean
Housing status*	**<0.001**	Farms with housed animals have a much higher mean prevalence
Whether animals have been moved	**0.004**	Farms with housed animals which have been moved during the previous 4 weeks have a lower mean prevalence
Whether animals have had a change in feed	**0.024**	Farms with housed animals which have had a change in feed during the previous 4 weeks have a lower mean prevalence
Whether the farm produces silage	**0.04**	Farms which have housed animals and which produce their own silage have a lower mean prevalence than other farms with housed animals
Whether the farm spreads manure on silage fields	**0.047**	Farms which have housed animals and which spread manure on their silage fields have a lower mean prevalence than other farms with housed animals
Whether the farm spreads slurry on silage fields	**0.027**	Farms with housed animals and which spread slurry on their silage fields have a lower mean prevalence than other farms with housed animals
Number of cattle	**0.012**	In housed groups, the presence of more cattle is associated with a lower mean prevalence
Whether deer are farmed	**0.036**	The presence of deer is associated with a higher mean prevalence; this is a poorly fitting factor
Water source for animals	**0.03**	Unhoused animals with water supplied from a natural source had lower prevalences than unhoused animals supplied from mains or private supplies

Factors with *P*-values <0.1 in the restricted multivariate analysis are given in bold type.
* Results for Housing status factor are for a marginal analysis.

only the number of finishing cattle was assessed in the multi-factor model. The factors defining whether a group had recently been subject to a move or changes in diet were partially aliased: most observations were of groups that were positive for neither or both factors. However, the results overall were consistent with both factors having a protective effect, but not cumulatively. A new factor was included in the multi-factor model, describing whether or not a group has been subject to either change. The factors summarising the use of manure or slurry in silage production were aliased with the factor defining home silage production. More detailed analysis suggested that the spreading of slurry on pasture was the key factor in this group. Terms involving the factor "source of animals" were found to have no statistically significant explanatory value in the multi-factor model and were removed. The results of the final multi-factor model are summarised in Table 5. Contrary to previous results (Synge et al., 2003),

Please cite this article in press as: Gunn, G.J. et al., An investigation of factors associated with the prevalence of ..., The Veterinary Journal (2007), doi:10.1016/j.tvjl.2007.08.024

Table 5
Results of the multi-factor analysis: levels of shedding in positive groups

Factor/variable	Effect	Log-odds ratio	SE	*P*-value
Animal Health Division	Scotland divided into three regions: Highlands; Central, Islands, North-East and South-East; and South West.Highlands exhibits a significantly higher mean prevalence than the portmanteau region	0.969	0.42	0.02
	The South West exhibits a significantly lower mean prevalence than the portmanteau region	−0.600	0.28	0.03
Housing status	Housed animals have higher mean prevalences	1.319	0.33	<0.001
Number of finishing cattle	Farms with >100 finishing cattle have significantly lower mean prevalences than those with <100	−0.702	0.23	0.004
Interaction between Housing status and 'Recent Changes in Housing or Diet'	Farms with housed animals and changes during the previous 4 weeks have higher mean prevalences than farms with unhoused animals. This effect is not formally significant	0.480	0.43	0.26
	Farms with housed animals and no changes during the previous 4 weeks have higher mean prevalences than farms with housed animals and recent changes	0.891	0.33	0.007
Water sourced from natural supply	Farms with animals at pasture have lower mean prevalences if the water is from a natural source	−0.708	0.35	0.04
Slurry spread on farm	Farms with housed animals which spread slurry on their silage fields have a lower mean prevalence than farms with housed animals which do not. This result is not formally statistically significant	−0.553	0.29	0.07

no consistent effects, statistically significant or not, were associated with the presence of geese on farms.

No statistically significant management system or temporal effects were observed in either the univariate or multi-factor model. In contrast, the restricted multi-factor analysis showed evidence of a geographical bias in prevalence, and this effect remained in the multi-factor model, unexplained by any of the explanatory factors. Hence, it can be concluded that the farm level prevalences do vary with region, in a fashion which is unaffected by the explanatory factors.

In both analyses, examination of the effect of field and laboratory operator confirmed that there was no evidence of any effect due to personnel bias.

Discussion

In a review of 26 published prevalence studies (Meyer-Broseta et al., 2001) the authors highlighted the problems caused by differing sampling and statistical methodologies. The microbiological methods used in this study were adopted as standard within the United Kingdom so the results may be compared with those obtained in England and Wales (Paiba et al., 2003) and subsequent studies, still to be published, carried out in Scotland under IPRAVE.

When the present study was designed, the use of 1 g faeces samples was considered to be more sensitive than rectal swabbing. The finding that VTEC O157 colonises lymphoid tissue at the recto-anal junction (Naylor et al., 2003) may explain why rectal sampling leads to higher prevalences at abattoirs (Chapman et al., 1997). With experimentally infected calves in the USA, enrichment culture of recto-anal mucosal swabs (RAMS) was found to be more sensitive than enrichment culture of 10 g faeces samples, once colonisation was established (Rice et al., 2003). These workers however pointed out that in transiently infected animals, i.e. animals shedding for <1 week, only faecal culture detected the organism. The estimates of prevalence described in this paper have been shown to underestimate by approximately a factor of two the prevalence within faecal pats as the organisms are not evenly distributed within the pats (Pearce et al., 2004). In this respect, the original objective of the study has not been fully delivered, since the prevalence estimate must be considered an underestimate. It is, however, useful as a lower bound, and is no more biased than other previously published results. Nevertheless, the analysis of risk factors presented is valid as an assessment of factors influencing (presumably) higher, and hence more detectable, levels of shedding in animals.

The estimate that 22.8% of groups of animals sampled contained at least one shedding animal fits closely with the findings of the longitudinal study in beef cows (Synge et al., 2003), when at least one shedding animal was found on 22% of 395 visits. That study clearly showed that repeat sampling increased the detection of VTEC O157, since after visiting each farm 12 times, the organism was detected in 28/32 (88%) herds.

In groups of cattle where shedding was detected, the level of observed shedding varied enormously (Fig. 2). Although >40% of these groups provided only one positive sample, 10% of the groups had all samples positive, suggesting that a high proportion of animals were shedding. It is plausible that in these herds at least some animals are shedding high numbers of organisms. This hypothesis was subsequently investigated by the IPRAVE project and the concept of super-shedders has been proposed (Matthews et al., 2005).

The main rationale for the analysis of the epidemiological dataset through two complementary analyses was to

ensure the goodness of fit of the associated statistical models. It is also reasonable to consider that the explanatory factors identified in the absence/presence analysis may be those most associated with the introduction and continued survival of infection on a farm. In addition those identified in the shedding levels analysis will be those which change the level of contact of animals with bacteria on the farm or which affect the propensity of carrier animals to shed. Obviously, the latter effects will change the propensity for infection surviving within the local farm population, and it is interesting that the significant factors identified in the two analyses are so disparate. A good example of this is the effect of herd size. The analysis demonstrated a statistically significant increase in the likelihood of groups being classed as positive if they were drawn from herds with more finishing cattle (Table 3). Interestingly, as reported in Table 5, farms with greater than 100 finishing cattle have statistically significantly lower mean proportions of animals shedding within the sample groups than farms with less than 100 finishing cattle.

There is no obvious epidemiological explanation of why the number of finishing cattle should have (apparently) opposing effects on the mean propensity to shed and the mean within-group prevalence. However, such an effect could arise from the interplay of the threshold properties of infection systems and the sampling scheme used in the prevalence study. It can be assumed that the between-animal infection rate of *E. coli* O157 within different farms is highly variable (perhaps explained by some of the risk factors listed in Table 5). The mathematical theory of epidemic systems (Anderson and May, 1991) would suggest that the probability of the infection dying out in small groups is higher than in larger groups. When a cross-sectional study is carried out, the sample of small groups will therefore tend to have proportionately more negatives. However, the samples from those small groups which are positive will disproportionately be drawn from those farms with high transmission rates, and hence with higher mean within-group prevalences.

In the univariate analysis, it was found that farms that purchased cattle for finishing rather than breeding their own replacements were more likely to present shedding animals. The effect of sourcing is however confounded with farm size, since larger farms are more likely to have bought in animals, and when the farm size is included in the model, source ceases to have any formal statistical significance. Detailed analysis of the effect of group size on the risk of shedding on open and closed farms separately indicates that group size is the important factor in determining risk.

The finding that a higher maximum age of cattle in the sampling group is associated with a lower risk of the group being positive is consistent with earlier work. For example, in one USA study, 0.2% adults and 0.65% weaned calves were found to be shedding (Hancock et al., 1994). Groups sampled on dairy farms with beef cattle are at a higher risk of being positive than those from other farms. Follow-up research into the pattern of infection on these units may yet help improve our understanding of the epidemiology. The suggestion that the presence of pigs on a farm is associated with a higher risk of shedding in cattle is contrary to previous findings (Synge et al., 2003). Pigs are not considered to be important in the epidemiology although clearly they can carry the organism (Borie et al., 1997; Chapman, 2000; Heuvelink et al., 1999). Contact between pigs and cattle on farms are, however, unlikely in most situations in Scotland, although indirect faecal contamination is probable. When evaluating this result, in particular, the risk of spurious false conclusions from multiple testing should be considered.

The spreading of slurry on grazing land was shown to increase the risk of groups of animals at pasture shedding VTEC O157. The spreading of manure on pasture was protective for housed groups. This is possibly because the majority of farms spread either slurry or manure on the pasture and it is known that the composting effects of dung heaps reduce the levels of bacteria in the faeces. The longitudinal study (Synge et al., 2003) identified wild geese as a risk factor for shedding in grazing cattle, but no such effect was observed in the present study.

The proportion of groups of cattle found to be shedding decreased significantly during in the study. The observed differences may reflect high year-to-year variability rather than a trend. There were no differences between the mean proportions of animals shedding within positive groups in different years. The suggestion that prevalence is declining is now being explored further in IPRAVE.

There was a cyclical effect with more herds shedding in the summer than the winter. This is in broad agreement with other studies. An early longitudinal study in a dairy herd in England showed two peaks of shedding, one in the early summer and one in November after housing (Mechie et al., 1997). A longitudinal study in cows (Synge et al., 2003) found the greatest number of groups shedding in the autumn months. Research in Aberdeen has demonstrated higher counts in the faeces of cattle in the summer months (Ogden et al., 2004). It is plausible that other, as yet unidentified, risk factors are influencing the group shedding risk in some months of the year. There was no variability in mean shedding proportions in positive groups between or within years that was not explained by other factors, predominantly housing.

The analysis showed that, while housed groups are no more likely to be shedding than grazing animals, the mean proportion of animals shedding in positive groups is statistically significantly greater in housed animals. This can perhaps be explained by the increased chance of bacterial transmission in housed animals or the greater chance of exposure from feed or water troughs (LeJeune et al., 2001). In the longitudinal study (Synge et al., 2003), groups of animals were more likely to be shedding when housed and an effect of bringing indoors was also noted. In the present study, housed animals that had had changes in diet or management in the previous four weeks showed higher mean shedding proportions than unhoused animals,

Please cite this article in press as: Gunn, G.J. et al., An investigation of factors associated with the prevalence of ..., The Veterinary Journal (2007), doi:10.1016/j.tvjl.2007.08.024

although this effect was not formally statistically significant. However, the mean proportion of shedding animals in housed groups which had not had such a recent change in diet or management was statistically significantly higher than in other housed groups. These results are consistent with a build up of exposure for housed animals from organisms colonising the environment.

Comparing groups of animals at pasture, cattle had lower mean shedding prevalences when they had access to a natural water supply. This may relate to the finding that water trough sediments in drinking troughs can be contaminated by cattle faeces and thence act as reservoirs for the spread of infection (LeJeune et al., 2001). Although cattle have occasionally been infected from natural watercourses, these results suggest that this is a low risk.

The study found no region of Scotland to be more or less likely to have shedding groups of cattle than any other, but within groups, the Highlands had significantly higher mean shedding prevalences and the South West significantly lower mean shedding levels compared to the rest of Scotland. These geographical differences could not be explained by the other explanatory variables included in the multi-factor model.

In addition to the determination of an estimate of the shedding prevalence of VTEC O157, the large size and design of this study has facilitated an extensive analysis of the risk factors affecting the shedding of the organism. The two-stage analysis of the data has proved successful in identifying risk factors that may influence different aspects of the epidemiology. In particular, it is interesting to note that housing status, although very important in affecting the mean prevalence within positive groups, shows no evidence of having any influence on whether or not a group is positive. It is therefore likely to operate as a risk factor purely at the within-group level. The interpretation of this housing factor is facilitated by the observed effects of recent changes in housing or diet, allowing us to infer some aspects of the likely infection dynamic as animals are brought in from pasture from a cross-sectional study.

Hypotheses about the likely transmission route of *E. coli* O157 will be informed by the findings about water supply and slurry and manure spreading as well as the identification of pigs as a risk factor, while the effect of cattle age and group size can be explained by reference to experimental studies and mathematical biology, respectively. In defining future research, the unexplained geographical and temporal variability might suggest alternative risk factors that vary in an unbalanced fashion across Scotland or across time. Perhaps most importantly of all, the apparent temporal decline in group prevalence, if continued into later years, would dramatically reduce the public health risk from *E. coli* O157 infection.

This study has produced a large volume of information. In general, however, because of the multiple hypothesis testing underlying these results, the results should be interpreted as indicating possible risk factors for further investigation, rather than as definitive statements of risk. Accordingly, these results have already guided researchers in several disciplines within the IPRAVE project in an effort to further elucidate the epidemiology of VTEC O157 carriage in cattle.

Uncited reference

Synge (1998). Q1

Acknowledgements

The Scottish Executive Environment and Rural Affairs Department (SEERAD) funded the execution (SAC/168/97) and the analysis (BSS/028/99) of this study. SEERAD Census Division provided a randomised list of farms from the sampling frame. Throughout the study, the guidance received from the project steering group was of great help. The co-operation from farmers who allowed us to sample their animals and who answered the questionnaire was greatly appreciated. The Wellcome Trust funded an International Partnership Research Award in Veterinary Epidemiology and staff employed in this project also contributed to the prevalence study.

References

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

Armitage, P., Berry, G., Matthews, J.N.S., 2002. Statistical Methods in Medical Research. Blackwell Science, Oxford, UK.

Borie, C., Monreal, Z., Guerrero, P., Sanchez, M.L., Martinez, J., Arellano, C., Prado, V., 1997. Prevalence and characterization of enterohaemorrhagic *Escherichia coli* isolated from healthy cattle and pigs slaughtered in Santiago, Chile. Archivos De Medicina Veterinaria 29, 205–212.

Chapman, P.A., 2000. Sources of *Escherichia coli* O157 and experiences over the past 15 years in Sheffield, UK. Journal of Applied Microbiology 88, 51S–60S.

Chapman, P.A., Siddons, C.A., Malo, A.T.C., Harkin, M.A., 1997. A 1-year study of *Escherichia coli* O157 in cattle, sheep, pigs and poultry. Epidemiology and Infection 119, 245–250.

Chapman, P.A., Wright, D.J., Norman, P., Fox, J., Crick, E., 1993. Cattle as a possible source of verocytotoxin-producing *Escherichia coli* O157 infections in man. Epidemiology and Infection 111, 439–447.

Chapman, P.A., Wright, D.J., Siddons, C.A., 1994. A comparison of immunomagnetic separation and direct culture for the isolation of verocytotoxin-producing *Escherichia coli* O157 from bovine feces. Journal of Medical Microbiology 40, 424–427.

Espie, E., Vaillant, V., Mariani-Kurkdjian, P., Grimont, F., Martin-Schaller, R., De Valk, H., Vernozy-Rozand, C., 2006. *Escherichia coli* O157 outbreak associated with fresh unpasteurized goats' cheese. Epidemiology and Infection 134, 143–146.

Foster, G., Hopkins, G.F., Gunn, G.J., Ternent, H.E., Thomson-Carter, F., Knight, H.I., Graham, D.J.L., Edge, V., Synge, B.A., 2003. A comparison of two pre-enrichment media prior to immunomagnetic separation for the isolation of *E. coli* O157 from bovine faeces. Journal of Applied Microbiology 95, 155–159.

Hancock, D.D., Besser, T.E., Kinsel, M.L., Tarr, P.I., Rice, D.H., Paros, M.G., 1994. The prevalence of *Escherichia coli* O157:H7 in dairy and beef cattle in Washington State. Epidemiology and Infection 113, 199–207.

Heuvelink, A., Schulten, S., Hoenderken, R., Bijker, P., deBoer, E., 1996. Verocytotoxin-producing *Escherichia coli* O157 in Dutch veal calves and adult cattle sampled at slaughterhouses. Tijdschrift Voor Diergeneeskunde 121, 642–646.

Heuvelink, A.E., Zwartkruis-Nahuis, J.T.M., van den Biggelaar, F., van Leeuwen, W.J., de Boer, E., 1999. Isolation and characterization of verocytotoxin-producing *Escherichia coli* O157 from slaughter pigs and poultry. International Journal of Food Microbiology 52, 67–75.

Karmali, M., 1989. Infection by verocytotoxin-producing *Escherichia coli*. Clinical Microbiology Reviews 2, 15–38.

Khakria, R., Duck, D., Lior, H., 1990. Extended phage-typing scheme for *Escherichia coli* O157:H7. Epidemiology and Infection 105, 511–520.

Lahti, E., Keskimaki, M., Rantala, L., Hyvonen, P., Siitonen, A., Honkanen-Buzalski, T., 2001. Occurrence of *Escherichia coli* O157 in Finnish cattle. Veterinary Microbiology 79, 239–251.

LeJeune, J.T., Besser, T.E., Hancock, D.D., 2001. Cattle water troughs as reservoirs of *Escherichia coli* O157. Applied and Environmental Microbiology 67, 3053–3057.

Louie, M., Deazavedo, J., Clarke, R., Borczyk, A., Lior, H., Richter, M., Brunton, J., 1994. Sequence heterogeneity of the eae gene and detection of verotoxin-producing *Escherichia coli* using serotype-specific primers. Epidemiology and Infection 112, 449–461.

Matthews, L., McKendrick, I.J., Ternent, H., Gunn, G.J., Synge, B.A., Woolhouse, M.E.J., 2005. Super-shedding cattle and the transmission dynamics of *Escherichia coli* O157. Epidemiology and Infection 134, 131–142.

Mechie, S.C., Chapman, P.A., Siddons, C.A., 1997. A fifteen month study of *Escherichia coli* O157:H7 in a dairy herd. Epidemiology and Infection 118, 17–25.

Meyer-Broseta, S., Bastian, S.N., Arne, P.D., Cerf, O., Sanaa, M., 2001. Review of epidemiological surveys on the prevalence of contamination of healthy cattle with *Escherichia coli* serogroup O157: H7. International Journal of Hygiene and Environmental Health 203, 347–361.

Møller Nielsen, E., Skov, M.N., Madsen, J.J., Lodal, J., Brøchner Jespersen, J., Baggesen, D.L., 2004. Verocytotoxin-producing *Escherichia coli* in wild birds and rodents in close proximity to farms. Applied and Environmental Microbiology 70, 6944–6947.

Naylor, S.W., Low, J.C., Besser, T.E., Mahajan, A., Gunn, G.J., Pearce, M.C., McKendrick, I.J., Smith, D.G.E., Gally, D.L., 2003. Lymphoid follicle-dense mucosa at the terminal rectum is the principal site of colonization of enterohemorrhagic *Escherichia coli* O157: H7 in the bovine host. Infection and Immunity 71, 1505–1512.

Ogden, I.D., MacRae, M., Strachan, N.J.C., 2004. Is the prevalence and shedding concentrations of *E. coli* O157 in beef cattle in Scotland seasonal? FEMS Microbiology Letters 233, 297–300.

Paiba, G.A., Wilesmith, J.W., Evans, S.J., Pascoe, S.J.S., Smith, R.P., Kidd, S.A., Ryan, J.B.M., McLaren, I.M., Chappell, S.A., Willshaw, G.A., Cheasty, T., French, N.P., Jones, T.W.H., Buchanan, H.F., Challoner, D.J., Colloff, A.D., Cranwell, M.P., Daniel, R.G., Davies, I.H., Duff, J.P., Hogg, R.A.T., Kirby, F.D., Millar, M.F., Monies, R.J., Nicholls, M.J., Payne, J.H., 2003. Prevalence of faecal excretion of verocytotoxigenic *Escherichia coli* O157 in cattle in England and Wales. Veterinary Record 153, 347–353.

Pearce, M.C., Fenlon, D., Low, J.C., Smith, A.W., Knight, H.I., Evans, J., Foster, G., Synge, B.A., Gunn, G.J., 2004. Distribution of *Escherichia coli* O157 in bovine fecal pats and its impact on estimates of the prevalence of fecal shedding. Applied and Environmental Microbiology 70, 5737–5743.

Pollard, D.K., Johnson, W.M., Lior, H., Tyler, S.D., Rozee, K.R., 1990. Rapid and specific detection of verotoxin genes in *Escherichia coli* by the polymerase chain reaction. Journal of Clinical Microbiology 28, 540–545.

Riley, L.W., Remis, R.S., Helgerson, S.D., McGee, H.B., Wells, J.G., David, B.R., Herbert, R.J., Olcott, E.S., Johnson, L.M., Hargrett, N.T., Blake, P.A., Cohen, M.L., 1983. Hemorrhagic colitis associated with a rare *Escherichia coli* serotype. New England Journal of Medicine 308, 681–685.

Rice, D.H., Sheng, H.Q.Q., Wynia, S.A., Hovde, C.J., 2003. Rectoanal mucosal swab culture is more sensitive than fecal culture and distinguishes *Escherichia coli* O157: H7-colonized cattle and those transiently shedding the same organism. Journal of Clinical Microbiology 41, 4924–4929.

Richards, M.S., Corkish, J.D., Sayers, A.R., McLaren, I.M., Evans, S.J., Wray, C., 1998. Studies of the presence of verocytotoxic *Escherichia coli* O157 in bovine faeces submitted for diagnostic purposes in England and Wales and on beef carcases in abattoirs in the United Kingdom. Epidemiology and Infection 120, 187–192.

Smith, H.R., Rowe, B., Adak, G.K., Reilly, W.J., 1998. Shiga toxin (verocytotoxin)-producing *Escherichia coli* in the United Kingdom. In: Kaper, J.B., O'Brien, A.D. (Eds.), *Escherichia coli* O157:H7 and other Shiga Toxin-Producing *E. coli* Strains. ASM Press, Washington, DC, pp. 49–58.

Synge, B.A., 1998. Methods used for the isolation and typing of *E. coli* isolates from livestock in Scotland. Concerted Action CT98-3935 verocytotoxigenic *E. coli* in Europe. Methods 1, 100–105.

Synge, B.A., Hopkins, G.F., 1996. Verocytotoxin-producing *E. coli* O157 – zoonotic implications. World Association for Buiatrics XIX Congress 2, 633–637.

Synge, B.A., Chase-Topping, M.E., Hopkins, G.F., McKendrick, I.J., Thomson-Carter, F., Gray, D., Rusbridge, S.M., Munro, F.I., Foster, G., Gunn, G.J., 2003. Factors influencing the shedding of verocytotoxin-producing *Escherichia coli* O157 by beef suckler cows. Epidemiology and Infection 130, 301–312.

Zhao, T., Doyle, M.P., Shere, J., Garber, L., 1995. Prevalence of enterohemorrhagic *Escherichia coli* O157-H7 in a survey of dairy herds. Applied and Environmental Microbiology 61, 1290–1293.

Please cite this article in press as: Gunn, G.J. et al., An investigation of factors associated with the prevalence of ..., The Veterinary Journal (2007), doi:10.1016/j.tvjl.2007.08.024