JOURNAL OF CLINICAL MICROBIOLOGY, May 2007, p. 1594–1603 Vol. 45, No. 5
0095-1137/07/$08.00+0 doi:10.1128/JCM.01690-06

Risk Factors for the Presence of High-Level Shedders of *Escherichia coli* O157 on Scottish Farms▽

Margo E. Chase-Topping,[1]* Iain J. McKendrick,[2] Michael C. Pearce,[3] Peter MacDonald,[4] Louise Matthews,[5] Jo Halliday,[6] Lesley Allison,[7] Dave Fenlon,[3] J. Christopher Low,[3] George Gunn,[3] and Mark E. J. Woolhouse[1]

Centre for Infectious Diseases, University of Edinburgh, King's Buildings, Edinburgh, United Kingdom[1]; Biomathematics & Statistics Scotland, University of Edinburgh, King's Buildings, Edinburgh, United Kingdom[2]; Scottish Agricultural College, Animal Health Group, Research Division, King's Buildings, Edinburgh, United Kingdom[3]; Department of Mathematics & Statistics, McMaster University, Ontario, Canada[4]; Institute for Comparative Medicine, Faculty of Veterinary Medicine, University of Glasgow, Glasgow, United Kingdom[5]; Centre for Tropical Veterinary Medicine, Easter Bush, Roslin, Midlothian, United Kingdom[6]; and Scottish E. coli O157 Reference Laboratory, Western General Hospital, Edinburgh, United Kingdom[7]

Received 16 August 2006/Returned for modification 3 November 2006/Accepted 2 March 2007

***Escherichia coli* O157 infections are the cause of sporadic or epidemic cases of often bloody diarrhea that can progress to hemolytic uremic syndrome (HUS), a systematic microvascular syndrome with predominately renal and neurological complications. HUS is responsible for most deaths associated with *E. coli* O157 infection. From March 2002 to February 2004, approximately 13,000 fecal pat samples from 481 farms with finishing/store cattle throughout Scotland were examined for the presence of *E. coli* O157. A total of 441 fecal pats from 91 farms tested positive for *E. coli* O157. From the positive samples, a point estimate for high-level shedders was identified using mixture distribution analysis on counts of *E. coli* O157. Models were developed based on the confidence interval surrounding this point estimate (high-level shedder, greater than 10^3 or greater than 10^4 CFU g^{-1} feces). The mean prevalence on high-level-shedding farms was higher than that on low-level-shedding farms. The presence of a high-level shedder on a farm was found to be associated with a high proportion of low-level shedding, consistent with the possibility of a higher level of transmission. Analysis of risk factors associated with the presence of a high-level shedder on a farm suggested the importance of the pathogen and individual host rather than the farm environment. The proportion of high-level shedders of phage 21/28 was higher than expected by chance. Management-related risk factors that were identified included the type of cattle (female breeding cattle) and cattle stress (movement and weaning), as opposed to environmental factors, such as water supply and feed.**

Verocytotoxin-producing *Escherichia coli* (VTEC), such as *E. coli* O157, is an important zoonotic agent with worldwide distribution. *E. coli* O157 may cause sporadic or epidemic cases of often bloody diarrhea that can progress to hemorrhagic colitis, thrombotic thrombocytopenic purpura, and hemolytic uremic syndrome (HUS) (21). HUS is a systematic microvascular syndrome that is initiated by secreted shiga toxins, with predominately renal and neurological complications, which are responsible for most deaths associated with *E. coli* O157 infection, particularly among elderly patients (22). Infection with *E. coli* O157 is a leading cause of acute renal failure in children (8). The incidence of *E. coli* O157 infection in Scotland is substantially higher than elsewhere in Great Britain (28) and abroad. Data from 2004 surveillance programs in Great Britain and the Republic of Ireland (Health Protection Surveillance Centre) have estimated the rate of infection per 100,000 population as 4.1, 1.35, 1.1, and 1.6 for Scotland, England and Wales, Northern Ireland, and the Republic of Ireland, respectively (28). Estimates from the U.S. FoodNet program are 0.9 per 100,000 population for 2004 (3).

Healthy cattle shed *E. coli* O157 in their feces (9, 17, 35), and this pathogen is present in most cattle operations (48). Cattle are the main reservoir host for *E. coli* O157 and other VTEC in the developed world (1) and play a significant role in the epidemiology of human infections (13). Outbreaks are attributed to consumption of contaminated food and water, animal contact, and person-to-person transmission (51). However, case control studies of sporadic infections, which account for the majority of cases of *E. coli* O157 infection in Scotland, have indicated direct contact with animals, their feces, and/or the farming environment as important risk factors (29, 40, 53). In addition, spatial analyses of the distribution of human cases have also identified an association between human infections and areas of high cattle density (20, 24, 36), providing further evidence of the risks associated with the farming environment. As such, considerable effort has been taken to determine the prevalence of *E. coli* O157 in cattle and identify possible risk factors for carriage. In the past decade, there have been two surveys conducted in Scotland to determine the prevalence and epidemiology of *E. coli* O157. The first study, funded by The Scottish Executive Environment and Rural Affairs Department (SEERAD), was conducted from March 1998 to May

* Corresponding author. Mailing address: Centre for Infectious Diseases, University of Edinburgh, Ashworth Laboratories, Kings Buildings, Edinburgh, Scotland EH9 3J5, United Kingdom. Phone: 44 (0)1316507263. E-mail: margo.chase@ed.ac.uk.

▽ Published ahead of print on 14 March 2007.

2000 (14) and involved 952 farms throughout Scotland. This remains the largest prevalence study for VTEC O157 ever carried out. The second survey, funded by the Wellcome Foundation International Partnership Research Award in Veterinary Epidemiology, is the source of data for this study.

Persistence and spreading of *E. coli* O157 within farms can be influenced by the duration, prevalence, and magnitude of shedding by individual animals and by bacterial survival and growth in the farm environment (4). Fecal shedding in individual cattle is transient (50), and *E. coli* O157 prevalence is known to have a highly skewed distribution (14), with the majority of cattle groups testing negative for the pathogen but a small proportion of animals shedding high numbers of *E. coli* O157.

In a recent publication, Matthews et al. (33) showed that the prevalence distribution of *E. coli* O157 is best explained when a small proportion of cattle are assumed to have much higher transmission rates than the others. These animals have been referred to as high-level shedders or "super shedders." Several other recent studies have suggested that some cattle may be high-level shedders of *E. coli* O157—that is, they harbor and shed the organism at much higher levels than other individuals (7, 30, 42). However, there appears to be no agreement in the literature as to the definition of a high-level shedder. High shedding has been defined as counts of *E. coli* that are $\geq 10^3$ (30) or $\geq 10^4$ CFU g^{-1} feces (37, 42, 41). In a recent study, Low et al. (30) found two subpopulations of animals—low-level shedders and high-level shedders, which were categorized by whether the *E. coli* levels were below or above 10^3 CFU g^{-1} feces, respectively. Low et al. (30) highlighted the possible importance of these high-level shedders to the epidemiology of *E. coli* O157 by establishing that the presence of high-level shedding is associated with a higher mean prevalence of low-level carriage on the farm, suggesting a higher transmission rate for high-level shedders. The occurrence of higher transmission rates was supported by mathematical models developed by Matthews et al. (32), who observed that a minority of high-level-shedding individuals were potentially responsible for most of the transmission, a common pattern observed in a variety of disease systems (55). Hence, significant reductions may be made by targeting the small portion of the population excreting large levels of *E. coli* O157, i.e., the high-level shedders (32). Therefore, the identification of risk factors influencing the presence of a high-level shedder are important to the development of interventions aimed at the prevention of high bacterial shedding and ultimately the development of an effective control strategy to reduce human disease resulting from *E. coli* O157 infections. The aim of this study was to identify farm-level risk factors for the presence of high-level shedders.

MATERIALS AND METHODS

Sampling design. Between March 2002 and February 2004, 481 farms were visited throughout Scotland as part of a cross-sectional survey to determine the prevalence of *E. coli* O157. These farms had been sampled previously for an earlier study commissioned by the SEERAD (March 1998 to May 2000), a larger survey involving 952 farms (14). The original farms were randomly selected from a list of 3,111 farms provided from the 1997 Scottish Agricultural and Horticultural Census data across all Scottish State Veterinary Service animal health divisions.

A stratified sampling plan was used to select farms to ensure that similar numbers were included from each region and that regions were sampled evenly over time. Cattle groups composed only of store (i.e., weaned cattle kept over winter on maintenance rations before finishing for slaughter) or finishing cattle closest to sale or slaughter were preferentially sampled. If such groups did not exist, one or more mixed groups with store or finishing cattle closest to sale or slaughter were sampled. From each group, fresh fecal pats were sampled. The number of fecal pats sampled in each group was determined from the number of cattle in the group, using a prescribed sampling schedule. It was assumed that, on average, 8% of the animals in positive groups would be shedding, with shedding distributed as seen in an earlier study commissioned by SEERAD (14). For each group, sufficient fresh pat samples were taken to ensure a mean 90% probability of detecting shedding of *E. coli* O157 if at least one shedding animal was indeed present. The size of the sample group on a farm was variable, ranging from 5 to 80 cattle with a median size of 27. Fecal samples were refrigerated at 5°C. The majority were refrigerated within 2 h of sampling, while a small number were held at ambient temperature before freezing on the day after sample collection.

Laboratory analysis. Within 48 h, fecal samples were examined by immunomagnetic separation (IMS) to detect the presence of *E. coli* O157 as described previously (44). Following IMS, one *E. coli* O157 isolate from each fecal sample was submitted to the Scottish *E. coli* O157 Reference Laboratory for phage typing (23) and tested for the presence of genes encoding the virulence factors verocytotoxin 1 (vtx_1), verocytotoxin 2 (vtx_2), and intimin (*eae*) using multiplex PCR (34, 52). Counts of *E. coli* O157 were estimated using the method described previously (44). One gram of feces from each *E. coli* O157-positive sample was suspended in 9 ml of maximum recovery diluent (Oxoid Ltd., Basingstoke, United Kingdom) and 0.1 ml of suspension spread onto each of two CT-SMAC plates. Plates were incubated at 42°C for 24 h. All typical non-sorbitol-fermenting colonies were counted, and a maximum of 10 such colonies were tested using anti-*E. coli* O157-coated latex reagent (Oxoid, Ltd.). The limit of accurate enumeration using this method was 100 CFU g^{-1} feces (44).

Farm management questionnaire. At the time of each farm visit, a farm management questionnaire was completed. The majority of questionnaires were administered by either of two operators. A detailed validation procedure was undertaken to validate the questionnaire and eliminate operator bias. Farm personnel were asked questions on cattle demography, health management practices, movement of animals, wildlife present, feed, and water supply. The questions were primarily closed, with a list of responses provided, although some questions offered the chance to respond as "other, please specify." A copy of the questionnaire is available upon request from the lead author.

Statistical analysis. (i) Mixture distribution analysis. High-level shedders were identified from the data using mixture distribution analysis. A mixture distribution is a composite of different statistical distributions arising from sampling a heterogeneous population, where it is not known a priori to which subpopulation any observation should be assigned. Similar methods have been used previously to identify high-level shedders in the classification used by Low et al. (30). MIX (Ichthus Data Systems, Hamilton, Ontario, Canada), implemented as package mixdist (5) for the R environment (47), was used to fit bimodal distributions to the *E. coli* O157 data and hence identify the high-level shedders within the population. Because the data ranged over several orders of magnitude, they were grouped into categories on a $\log_{10}$ scale. False negatives are relatively frequent when the true mean is less than 10^2 CFU g^{-1}, so observations below this threshold (which would represent low-level carriage) were excluded to avoid biasing the fitting procedure. Parameters of the mixture distribution (mean, standard deviation, and proportion) were calculated as maximum-likelihood estimates using a combination of the expectation-maximization (EM) algorithm and Newton-type methods. Models were fitted assuming two normal components, two gamma components, a normal component with a two-parameter gamma, and a normal component with a three-parameter gamma. The current release of the mixdist package does not fit normal/gamma mixtures; this was done with a special-purpose R function available from the authors.

The goodness of fit of all models was assessed. The normal/three-parameter gamma model gave rise to the best fit, but it was not statistically significantly better than the fit from the other models. The fit of all of the models was acceptable, so the normal/normal was preferred as making fewer assumptions about the properties of the data, since there is no extensive data set available to test such assumptions.

The point at which an observation would be classed a posteriori as equally likely to have arisen from either subpopulation, given no prior belief about the relative numbers in the subpopulations, was chosen as the point estimate of the threshold defining high-level shedders. As a measure of the variability of the estimate, 95% confidence estimates were generated for the threshold. Using the computer parameter estimates and their variance-covariance matrix and assuming joint asymptotic normality of the maximum-likelihood estimates, a distribu-

tion of thresholds was generated by repeated simulation of pseudoestimates of the parameters from a multivariate normal distribution and calculation of the associated pseudothresholds. The 2.5% and 97.5% points of the resulting distribution were used to define a confidence interval (CI) for the estimated threshold. These 95% confidence estimates were used as the basis of all subsequent statistical analyses.

(ii) Single-variate and multiple-variate analysis. Univariate analysis on the virulence factors, phage type, and management categorical variables was performed using either the chi-square test or Fisher's exact test, as appropriate. StatXact v.5 (Cytel Software Corp., Cambridge, MA) was used to evaluate the significance of contingency tables. Analysis of the continuous variables (e.g., total number of cattle) was performed using either analysis of variance or Mann-Whitney test, depending on the structure of the data. Prior to analysis, it was specified that results with P values of <0.05 would be reported as exhibiting formal statistical significance.

Analyses using multiple explanatory variates were performed using generalized linear mixed models (GLMM) (Proc Glimmix SAS, version 9.3.1; SAS Institute Inc., Cary, NC) and generalized linear models (GLMs) (Proc Logistic/Proc Genmod SAS, version 9.3.1) to (i) examine the hypothesis regarding the presence of a high-level shedder as a risk factor for the observation of low shedders (30) and (ii) examine the associations between the presence of high levels of *E. coli* O157 in feces and management practices, conditional on shedding being present on the farm. Only positive farms were used in the multiple variable analyses.

A GLMM assuming a binomial response term and a logit link was fitted to describe the number of low shedders relative to the number of non-high-level shedders. The presence of at least one high-level shedder was fitted as a risk factor (coded 0 versus 1 for presence/absence). Hence, any direct relationship between response and explanatory variable was removed. The farm was fitted as the sole random effect to model overdispersion in the data. The presence of a high-level shedder was fitted as a fixed effect.

Associations between the observation of *E. coli* O157 in feces and management practices were determined using GLMs assuming a binomial response and a logit link. The response variate was a binary variable corresponding to the presence or absence of a high-level shedder on each farm. Prior to building the model, univariate associations between *E. coli* O157 and each of the management variables were performed as a screening step. Factors that were associated significantly with the outcome variable with a P value of ≤ 0.25 were considered for further analysis. These were included as candidate variables in a logistic-regression model (Proc Logistic/Proc Genmod SAS, version 9.3.1). Three factors had a prevalence of zero in at least one category and could not be analyzed in the logistic-regression model because of the resulting failure to fit. These factors were calving, spring calving, and spreading of human sewage. The factors calving and spring calving were identified as exhibiting statistically significant effects by Fisher's exact test ($P = 0.02$ and $P = 0.001$, respectively), where the probability of observing a high-shedding animal was higher with either of these factors present. The variable "calve-month" was created to examine the intensity of calving. One factor category had to be removed from each of the following factors as a result of insufficient data in the category: sample year, interviewer, management type, and type of spread. None of these factors was significant in the final models. Region and season were forced into the model as design factors. Seasons were defined as winter, comprising December, January, and February; spring, comprising March, April, and May; summer, comprising June, July, and August; and autumn, comprising September, October, and November. Three regions were defined: 1 = North Scotland (including Highlands, Islands, and North East); 2 = Central Scotland; 3 = South Scotland. A hierarchical forward selection and a backward elimination approach with swapping (reassessment of previously included or excluded variables) were used. The change in the deviance of the model was monitored as an indicator of the improved fit of a model. Variables were added and removed based on significant improvement in the mean deviance after changes to the model. Two-way interactions were also tested in this manner. For the final model, the Hosmer-Lemeshow goodness-of-fit statistic was computed (19).

To check for multicollinearity between factors in the final model, correlations were examined for binary and nominal variables. In addition, the stability of the model was checked by systematic removal of variables. Diagnostics were performed and plots of residuals were examined, confirming the goodness of fit of the model. Odds ratios and their associated 95% CIs were estimated in the final logistic model for factors statistically significantly associated with the presence of *E. coli* O157-high-shedding animals.

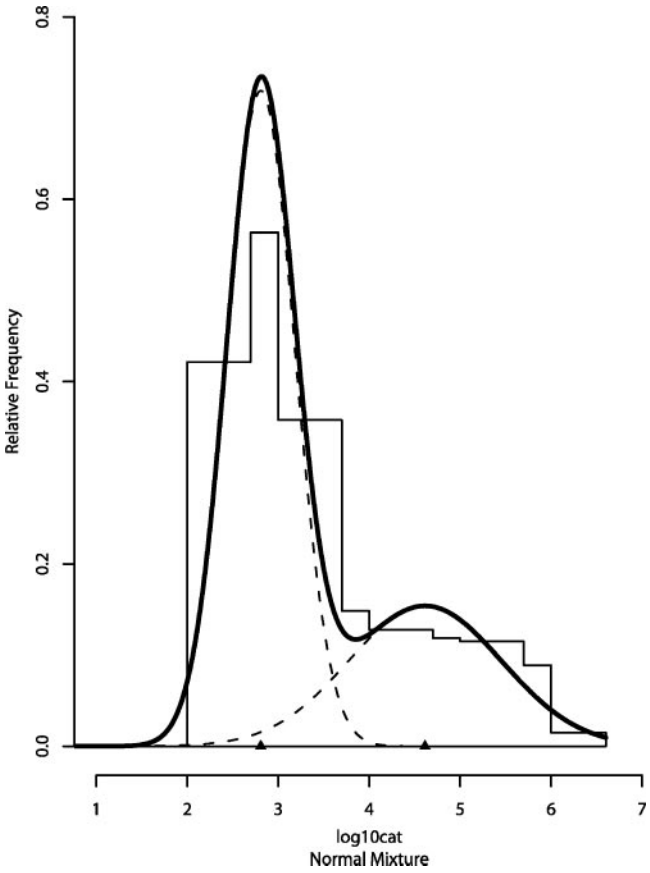

FIG. 1. Results of mixture distribution analysis. Normal-normal distribution. Point estimate of threshold for high-level shedding, 3,135; 95% CI for threshold estimate, 1,658 and 10,395.

RESULTS

Mixture distribution analysis. The concentration of *E. coli* O157 in fecal pats varied both within and among farms. Counts ranged from <100 to $>36 \times 10^6$ CFU g^{-1}, with the majority of farms (74%) only being detected by IMS as a result of having individual concentrations of <100 CFU g^{-1}, the limit of accurate enumeration (44). Since the majority of the data (74%) were below the limit of accurate counting, these values had to be set completely inside the first subpopulation of the mixture distribution. Since the means and standard deviations for the first component would not be reliable in this situation, such observations were removed from the analyses and MIX was run with the remaining data. Figure 1 illustrates the fit of the preferred model to the observed data. The parameters mean, standard deviation, and population proportion of the first and second subpopulations of the distribution are as follows: subpopulation 1, 2.808, 0.374, and 0.675; subpopulation 2, 4.613, 0.842, and 0.325. The estimated threshold for a high-level shedder was chosen as the point of overlap of the two unscaled distributions, approximately 3,135 CFU g^{-1}. Note that this does not match the value which might be inferred from Fig. 1, where the components are scaled relative to the sizes of the observed populations. Confidence bands of 1,658 CFU g^{-1} to 10,395 CFU g^{-1} feces were generated around this estimate, approximated as $\geq 10^3$ and $\geq 10^4$ CFU g^{-1} feces, and used to

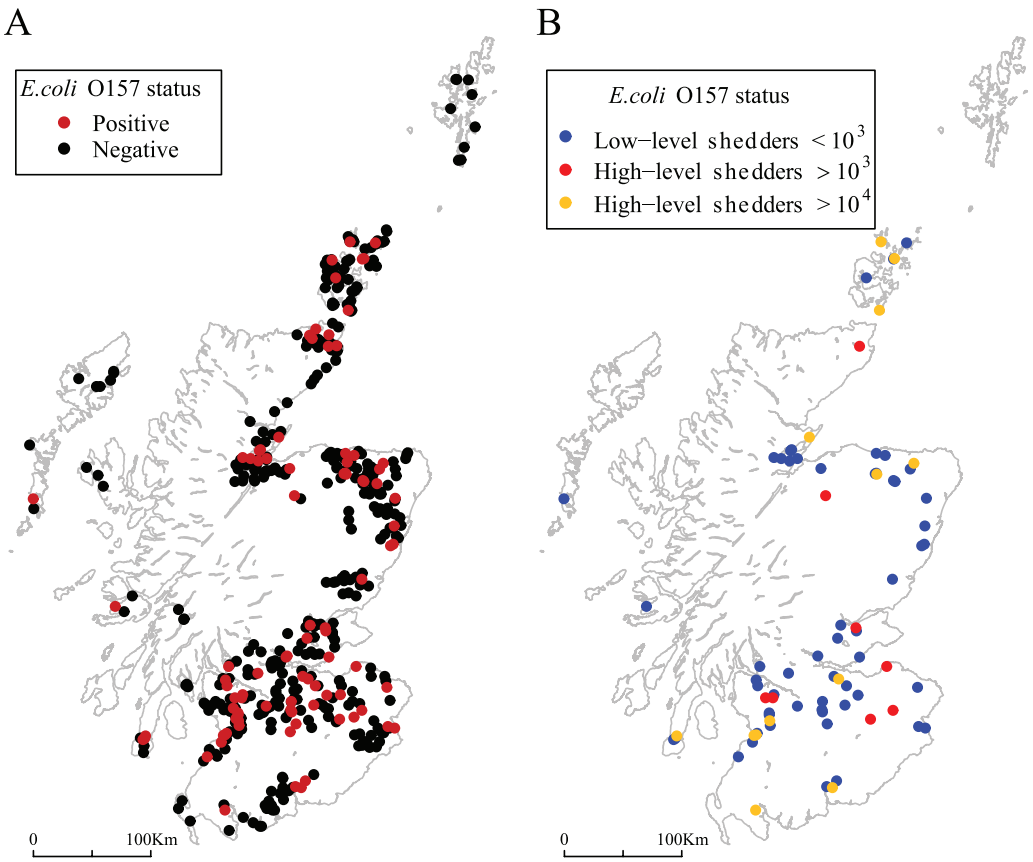

FIG. 2. Geographical distribution of farms. (A) Distribution of all 481 farms sampled in Scotland between March 2002 and February 2004. Red dots indicate positive farms ($n = 91$). Prevalence estimated at 18.9% (exact binomial 95% CI = 15.5 to 22.7). (B) Distribution of positive farms used in all analyses ($n = 77$) showing the designation of shedding. The maps were created on the basis of border data from http://www.edina.ac.uk.

define high-shedding model I and high-shedding model II. For analysis, a farm was considered to contain a high-level shedder if at least one sample from that farm had *E. coli* counts at concentrations of $\geq 10^3$ CFU g^{-1} or $\geq 10^4$ CFU g^{-1}. This yielded 21 or 13 high-level shedding farms, respectively.

General epidemiology. Fig. 2 is a map of Scotland with the locations of all the farms visited in this study and their status with respect to the presence/absence of *E. coli* O157. Ninety-one farms were positive for *E. coli* O157 (18.9%; exact binomial 95% CI, 15.5 to 22.7) (Fig. 2A). Of the positive farms, there were 14 farms which did not present a complete set of count data. These farms were designated unknown, since they could not be classified, and were excluded from any subsequent analyses. Of the remaining positive farms ($n = 77$), 21 had a concentration of *E. coli* of $\geq 10^3$ CFU g^{-1}, 13 of which had a concentration of *E. coli* of $\geq 10^4$ CFU g^{-1} (Fig. 2B).

Risk factor analysis. (i) Associations with virulence factors. Positive isolates were characterized by determining the phage type (PT) and examining for the presence of genes encoding intimin (*eae*) or verotoxins (vtx_1 or vtx_2). With the exception of phage type, no associations were found between high shedding and these virulence factors. All but one isolate carried the *eae* gene encoding intimin, 98% carried the vtx_2 verocytotoxin gene, and only one isolate carried neither the vtx_1 gene nor the vtx_2 gene. There was no statistical evidence of association between the presence of verototoxins and high shedding ($P > 0.05$). There were 11 different PTs identified in this study. The main PTs identified were 21/28, 32, and 8, which comprised approximately 46%, 19%, and 12% of the samples, respectively. The predominant PTs were examined to determine if there was any association between a specific PT and high shedding. The null hypothesis was that the proportion of PT 21/28 was the same for high-level and low-level shedders. High-level shedders were more likely to be PT 21/28 than low-level shedders ($\geq 10^3$ CFU g^{-1}, chi square = 9.37; $P = 0.003$; exact odds ratio = 2.35; $\geq 10^4$ CFU g^{-1}, chi square = 6.56; $P = 0.02$; exact odds ratio = 2.90) (Fig. 3). The proportion of PT 32, however, is lower than one might expect from chance alone among high-level shedders compared that among with low-level shedders, but only for shedding levels of $\geq 10^3$ CFU g^{-1} ($\geq 10^3$ CFU g^{-1}, chi square = 11.37; $P = 0.001$; exact odds ratio = 0.184; $\geq 10^4$ CFU g^{-1}, chi square = 2.48; $P = 0.133$) (Fig. 3). No other associations were statistically significant.

(ii) Associations with higher farm prevalence. Low et al. (30) determined that the presence of high-level carriage was statistically significantly associated with a higher risk of low-level carriage. Figure 4 shows the distribution of prevalence on high-level shedding and low-level shedding farms in this study. The graph shows that a high mean prevalence of low-level carriage occurs more frequently on high-level shedding farms, with the majority of low-level shedding farms having prevalences in the range 0 to 0.1. The associated GLMM model was

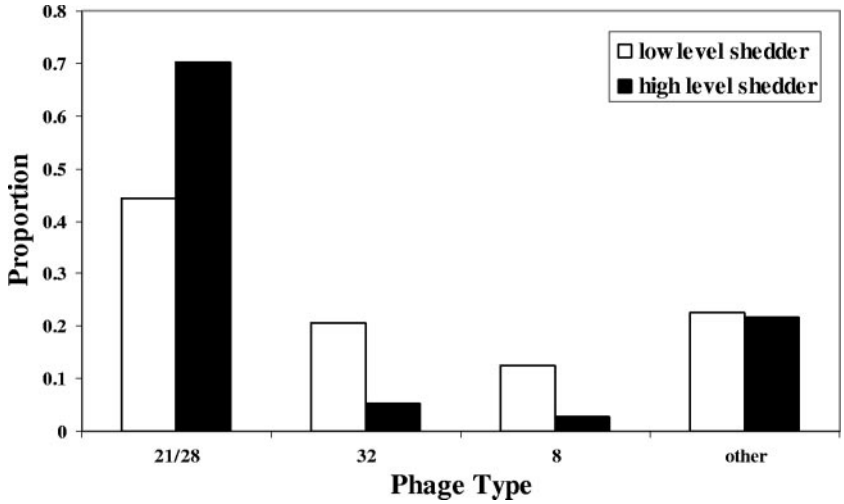

FIG. 3. Comparison of proportions of the three major phage types (and others) for low-level and high-level shedders using the point estimate of the high-level threshold (≥3,135). Chi square was significant (chi square = 9.135; $P = 0.003$; exact odds ratio = 2.97).

statistically significant ($\geq 10^3$ CFU g^{-1}, $F = 7.66$; df = 1,74; $P = 0.007$; odds ratio = 2.36 [1.28 to 4.35]; $\geq 10^4$ CFU g^{-1}, $F = 7.74$; df = 1,69; $P = 0.007$; odds ratio = 2.74 [1.35 to 5.58]). Thus, the observation of a high-level shedding animal on a farm more than doubled the odds of sampling low-level shedders in the local population.

(iii) Associations with farm management. Univariable analysis resulted in 25 variables remaining for multiple-variable analysis. Final logistic models for the two definitions of high-level shedders are shown in Table 1. A maximum of five variables remained in the final models. No interactions were statistically significant ($P > 0.05$). The number of positive isolates was the most statistically significant variable and was entered first into both models. Significant risk factors identified in model I (high-level shedder, $\geq 10^3$ CFU g^{-1}) include the observation of a large number of positive isolates, bringing on (BO) female breeding cattle as opposed to BO fattening cattle (protective), calving for fewer than 6 months of the year (number of calving months), and the use of straw for feed. Significant risk factors identified in model II (high-level shedder, $\geq 10^4$ CFU g^{-1}) include the observation of large number of positive isolates and the presence of cows in the sample group. The Hosmer-Lemeshow goodness-of-fit statistics were 3.0146 ($P = 0.88$) and 2.1729 ($P = 0.95$) for model I and model II, respectively, providing no evidence of a lack of fit in the selected models (19).

(iv) Relationship with number of positives. In both model I and model II, the most significant risk factor was the variable "number of positives," defining the number of positive pats identified on the farm. The statistical significance of this variable may reflect an association between the presence of a high-level shedder and higher levels of transmission, or it may be a spurious result, where the observation of a larger sample of positive pats defines a larger pool of samples, any of which might be observed as being from a high-level shedder on a purely random basis. If a constant fraction of positive pats is

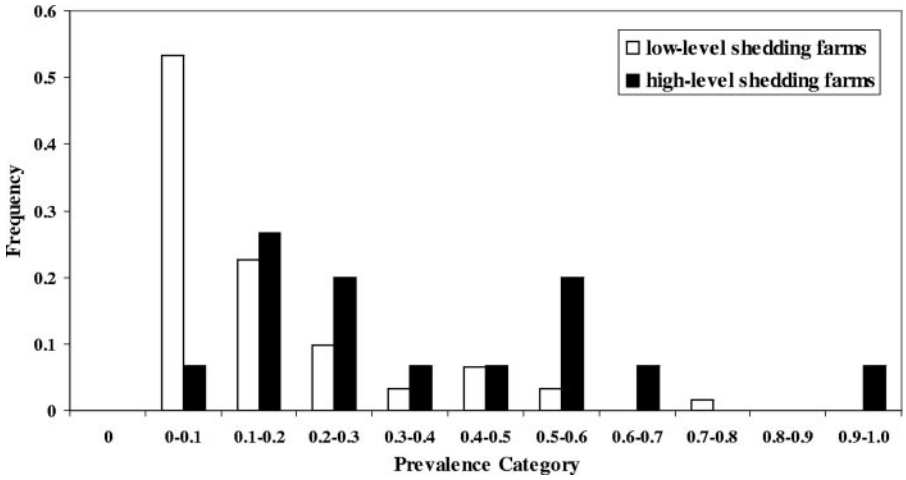

FIG. 4. Comparison of distribution of prevalences of *E. coli* O157 on low-level- and high-level-shedding farms using the point estimate of the high-level threshold (≥3,135). The resulting GLMM model was significant ($F = 5.27$; df = 1,71; $P = 0.025$; odds ratio = 2.28 [1.11 to 4.65]).

TABLE 1. Odds ratios and 95% CIs for estimated effects in final generalized linear models[a]

Model and variable	Estimated effect (SE)	Odds ratio	95% CI	*P* value
Model I (high-level shedder; $\geq 10^3$ CFU g^{-1} feces)				
Region 1	0.97 (1.368)	9.37	0.181–38.5	0.22
Region 2	1.31 (1.174)	3.70	0.371–36.9	0.26
Region 3 (baseline)	0.00 (0.000)			
Season 1	6.11 (2.746)	451	2.07–98200	0.03
Season 2	4.12 (1.865)	61.4	1.59–2380	0.03
Season 3	2.85 (1.581)	17.2	0.777–382	0.07
Season 4 (baseline)	0.00 (0.000)			
No. positive	4.26 (1.524)	70.8	3.58–1400	0.005
BO female cattle	4.12 (1.512)	61.8	3.19–1200	0.006
No. of calving mo	−5.48 (2.088)	0.004	<0.001–0.251	0.009
BO fattening cattle	−4.81 (2.033)	0.008	<0.001–0.436	0.02
Straw	3.19 (1.557)	24.3	1.15–513	0.04
Model II (high-level shedder; $\geq 10^4$ CFU g^{-1} feces)				
Region 1	1.88 (2.507)	6.53	0.048–890	0.23
Region 2	−0.34 (1.131)	0.713	0.078–6.57	0.76
Region 3 (baseline)	0.00 (0.000)			
Season 1	2.86 (1.605)	17.5	0.755–408	0.07
Season 2	2.90 (1.582)	18.2	0.821–405	0.07
Season 3	6.81 (2.629)	909	5.26–157000	0.01
Season 4 (baseline)	0.00 (0.000)			
No. positive	4.21 (1.587)	67.6	3.01–1520	0.008
Cows in sample group	5.06 (1.862)	26.3	4.09–6040	0.007

[a] Season and region are included as design variables but were not significant in either model I or model II. Odds ratio = exp(estimated effect); 95% CI = exp(estimated effect ± 1.96 SE).

produced by high-shedding animals, the chance of identifying a high-level shedder will increase with the number of positives isolated. In an attempt to ascertain the true meaning of the observed statistical significance, the GLM model described above was refitted with "number of high-level pats" from "number of positive samples" as the binomial response variable and the same explanatory variables with the exception of the "number of positives" variable, which was excluded. The parameter estimates from this model were used in a predictive model to generate pseudorandom variables, which were summarized as binary responses (absence/presence of high-level pats). Many (10,000) models were fitted to these pseudo-data sets using the same explanatory variables (now including the "number of positives" variable) in order to obtain a distribution of the coefficients for the "number of positives" variable under the null hypothesis that the variable "number of positives" is not directly a risk factor for the presence of high shedders. The coefficients from the original GLM (Table 1) were compared to this distribution. If the original estimated parameter was within the 95% modal zone of the distribution, then the observed "number of positives" regression coefficient was consistent with the effect being purely an artifact of the sample size. However, if the original parameter was in the tails of the simulated distribution, then the significance of the "number of positives" variable was indicative of a direct association between the presence of a high-level shedder and higher transmission. The resulting graphs for the two levels of high-level shedding ($\geq 10^3$ CFU g^{-1} and $\geq 10^4$ CFU g^{-1}) are shown in Fig. 5. The locations of the original regression coefficients for the "number of positives" variable were not within the statistically significant tails of the distributions (two-tailed test, $\geq 10^3$ CFU g^{-1}, $P = 0.34$; $\geq 10^4$ CFU g^{-1}, $P = 0.15$).

DISCUSSION

Recent literature on *E. coli* O157 has highlighted the importance of the high-level shedder or super shedder as a possible target for control on farms (26, 33). The concept of a high-level shedder/super shedder/super spreader was highlighted in a recent outbreak of severe acute respiratory syndrome (25) and more generally (10, 26) for situations in which a few individuals were responsible for disproportionate numbers of transmission events. Modeling work by Matthews et al. (33) describing the distribution of *E. coli* O157 counts suggested that the observed distribution may reflect epidemiologically important between-animal variation in shedding levels, where a small proportion of cattle are assumed to have much higher transmission rates than the others. Variable levels of *E. coli* O157 shedding have been observed in experimental studies (37), observational studies (45, 30), and the results of this research. Low et al. (30) observed different subpopulations of animals, defined by the level of shedding of *E. coli* O157. The mixture distribution analysis performed in this study also suggested that there was more than one population of animals shedding *E. coli* O157, with the point estimate of high-level shedding located at approximately 3,135 CFU g^{-1} feces (low- and high-confidence thresholds, $\geq 10^3$ and $\geq 10^4$). It is possible that all shedding reaches high levels at some point in the infection process. However, a recently published longitudinal study found evidence of persistent high shedding ($>10^3$ CFU g^{-1}) (45). In

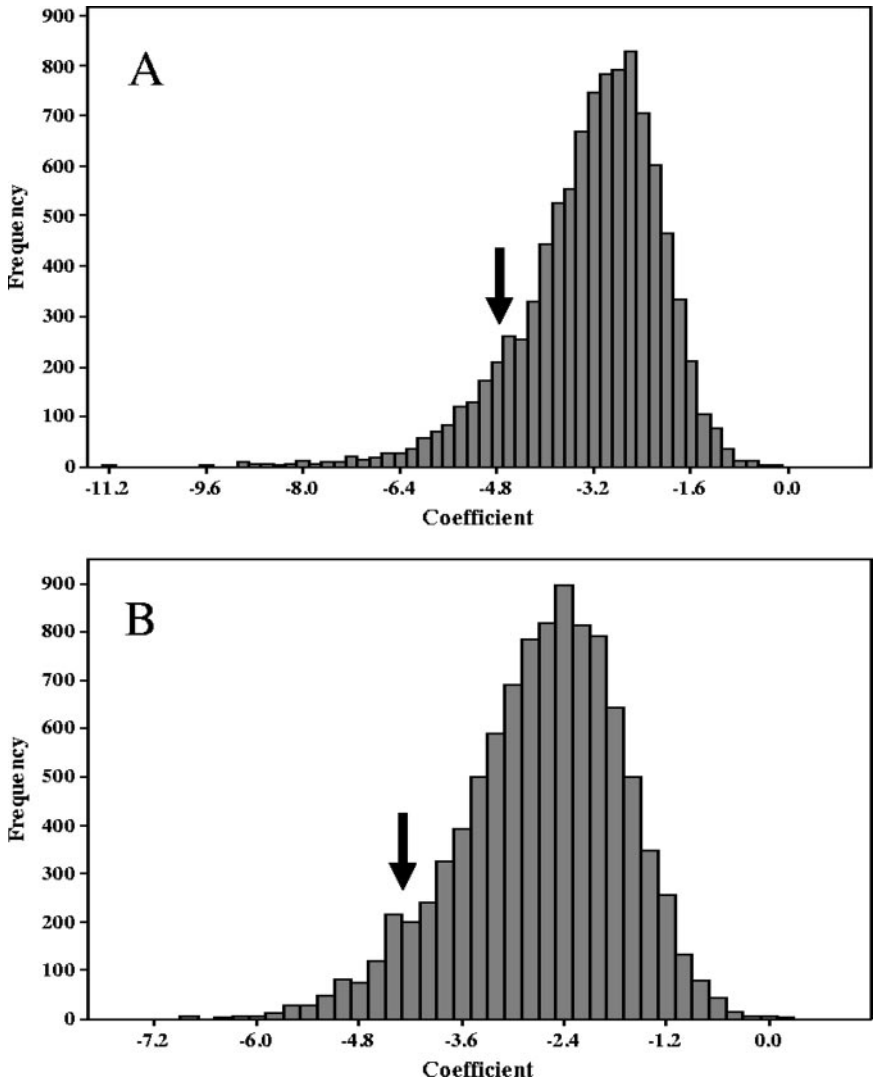

FIG. 5. Distribution of regression coefficients for "number of positives" variable from GLM analysis. Graphs summarize 10,000 runs. The arrow represents the location of the coefficient from the original GLM. (A) Model using $\geq 10^3$ CFU g^{-1} feces to define high-level shedder. The original parameter (−4.26) was not extreme relative to the pseudodistribution (two-tailed test, $P = 0.34$). (B) Model using $\geq 10^4$ CFU g^{-1} feces to define high-level shedder. The original parameter (−4.21) was not extreme relative to the pseudodistribution (two-tailed test, $P = 0.15$).

addition, experimental work has failed in some instances to produce infections with high levels of shedding (David Gally, personal communication). As such, there is evidence of the existence of animals that shed (and possibly transmit) *E. coli* O157 at high levels.

At present, it is unknown what factors lead to the generation of high-level shedders (e.g., cattle genetics or management, bacterial strain, or gut flora). However, experimental infections have shown that only cattle colonized at the terminal rectum excrete high levels of *E. coli* O157 in the feces, as opposed to those animals that are simply carrying *E. coli* O157 but without true colonization of the mucosal epithelium (30, 37). Understanding such colonization is a crucial step in the development of methods for control. Analyses of risk factors associated with the presence of high-level shedders suggest that the pathogen and the individual host are more influential for transmission than the environment. The proportion of high-level shedders of PT 21/28 was higher than expected by chance, suggesting that this subgroup of *E. coli* O157 is important. PT 21/28 was the most common phage type identified in this study (46%), and this correlates with Scottish human infection data (58% of Scottish *E. coli* O157 human infections were identified as PT 21/28 in 2004 (27). PT 21/28 is of particular concern because of its association with more-severe human morbidity. In the United Kingdom and Ireland (1997 to 2001), the risk of developing diarrhea-associated HUS was significantly higher for children in Scotland infected with PT 21/28 than with other phage types (31). In Scotland from 1997 to 2001, 61% of HUS cases in individuals less than 16 years old were attributable to infection with PT 21/28. The odds of shedding phage type 21/28 were more than two times higher for high-level shedders than for low-level shedders. An association between higher counts of *E. coli* O157 and PT 21/28 was supported by a recently published study (15). Halliday et al. (15) have also identified specific risk factors for the presence of PT 21/28.

Although phage type was statistically significant in this study and the study by Halliday et al. (15), it does not fully determine the observed bacterial count, and its functional significance is unknown. Phage typing is a widely used method for subtyping of *E. coli* O157 isolates, based upon the reactions of different isolates to a panel of bacteriophages. Sensitivity/resistance to different bacteriophages is a phenotypic expression. It is unknown whether or not differences in phage type reflect genotypic variation. Many factors determine whether a particular phage will lyse a host strain, including physical attributes, such as surface receptors on the organism, and genetic attributes, such as lysogeny with related phages. In *Salmonella*, for example, the phage type changes in response to possession of particular plasmids and stress, leading to changes in the expression of surface antigens (Derek Brown, Salmonella Reference laboratory, personal communication). In addition, there are differences in the abilities of different phage types to infect/colonize different hosts. It appears, however, that PT 21/28 may be a marker for some genetic difference or altered expression in *E. coli* O157, resulting in that particular subgroup being shed from cattle at higher levels. It has been suggested recently that the altered regulation of the type III secretion system in PT 21/28 strains compared to other phage type strains may enable the bacteria to colonize and be excreted at higher levels (6, 38). It is known that type III secretion is essential for bovine colonization, which in turn increases the likelihood of human infection (46).

PT 21/28 has been the most common phage type in Scotland since 1996/1997, when it overtook PT2. However, infection with PT 21/28 now appears to be declining in humans (27) and possibly cattle (unpublished data). PT 21/28 has been a particularly successful *E. coli* O157 subgroup. Analysis of the pulsed-field gel electrophoresis profiles of different PT 21/28 isolates in humans has revealed "recurrent clones" that have persisted since 1994, an indication of the organism's ability to establish a successful relationship with its host, allowing it to persist and propagate. The high prevalence and transmission of PT 21/28 could be associated with the higher propensity to super shedding identified in this study, although they could also be related to persistence outside the host environment.

The statistically significant risk factors identified from the wide range of managerial and environmental factors recorded also indicate the importance of the individual host rather than the environment in high-level shedding. No factor relating to the environment, water, or food was identified as a statistically significant risk factor. Those identified as risk factors related to the type of cattle on a farm (e.g., female breeding cattle) and cattle stress (movement and weaning).

The association between cattle movement and the presence of *E. coli* O157 is expected. It is known that movements of cattle are critical to the transmission of other infectious livestock diseases, including bovine tuberculosis, bovine spongiform encephalopathy, and foot and mouth disease (12). There is much evidence in the literature to support the association between the movement of cattle and the risk of *E. coli* O157 shedding (16, 39, 49, 54). However, the results of this study suggest that for high-level shedders to be present, it is not cattle movements per se but the type of cattle that are moved. Movement of female breeding cattle was a risk factor in this study, yet the movement of fattening cattle was protective. Farms with cows, young animals, and concentrated calving periods all appear to be at more risk of observing a high-level shedder. Farms with high-level shedders tended to calve in concentrated periods of 1 to 6 months' duration (81%), generally in March/April, whereas there was a more even distribution of calving among low-level-shedding farms. Spring calving was identified as a strong risk factor in the univariate analysis ($P = 0.001$). Calves born in the spring are weaned in autumn at the time of housing. From weaning until 12 months of age, calves are believed to be a major risk group for shedding VTEC O157 (18, 39, 43). Higher prevalences have been detected for calves after weaning (11, 16, 56). Studies conducted in other countries have reported similar *E. coli* O157 excretion rates for calves but with higher prevalences for weaned calves than for unweaned calves (11). Others did not find this (4). It is possible that young animals not only may have a higher prevalence of *E. coli* O157 (2) but also may shed higher numbers. The increase in prevalence following weaning may be associated with the loss of protective factors present in maternal milk or dietary changes with secondary changes to the gut microflora, as well as increased contact with other calves and, potentially, weaning stress (39). These arguments, however, do not explain why spring calving should be a particular risk. Previous studies (14) have identified the housing of animals as a risk factor for all levels of *E. coli* O157 shedding. It is possible that the higher risk of observing high shedders with spring calving arises from the effect of the weaning risk being temporally confounded with the housing risk.

The current interest in the identification of the presence of high-level shedders on a farm arises from the possible link between high-level shedding and higher levels of transmission. Recent modeling work (32, 33) has indicated that data collected from shedding animals is highly consistent with this possibility. Unfortunately, the results of the current study, although suggestive, cannot establish such an association. In the risk factor analysis, the number of positive animals on a farm was the most significant predictor of the presence of a high-level shedder. In addition, an association was observed between the presence of a high-level shedder and a high proportion of low shedders, similar to that observed by Low et al. (30). These results are consistent with higher transmission rates for high-level shedding animals but could arise from sampling bias or via the random occurrence of high-level shedders among positive animals whether or not they have higher transmission rates. In an attempt to shed more light on the situation, further statistical analysis was performed. The results suggested that the association between the observation of high-level shedders and a higher number of positive fecal pats on the farm could be explained as a function of the sample size without requiring any additional assumption of higher transmission rates. Obviously this result does not invalidate the transmission hypothesis, which is itself consistent with an association between numbers of high shedders and numbers of low-shedding animals. Although not statistically significant, the regression coefficient for the original fitted model was in the tail of the estimated null distribution, suggesting that there may be a genuine effect, although we did not have sufficient data to statistically confirm any association. It is possible that such a relationship applies only to PT 21/28 shedding, in which case the inclusion of other phage types in the data would

weaken the power of the statistical exercise. Unfortunately, there were insufficient PT 21/28 data to test this theory. In addition, we cannot exclude the hypothesis that there are hidden confounding variables which increase the prevalences of both low shedders and high shedders, perhaps to different extents. However, it is impossible to test such a theory with survey-based data.

This paper has identified risk factors for high-level shedders which indicate the importance of the pathogen and individual host rather than the environment. The type of cattle on the farm, especially the presence of young female breeding cattle, was identified as a risk factor for high-level shedding of *E. coli* O157. In addition, there are significant associations with the strain type. High-level shedders were more likely to be shedding PT 21/28. This research highlights the importance of heterogeneity in the population and the complexity of the problem of elucidating the *E. coli* O157 problem in Scotland and abroad. It is our hope that these results will help in the development of interventions to reduce or prevent shedding from cattle (and therefore protect human health) and will help generate models that predict the efficacy of such interventions.

ACKNOWLEDGMENTS

We thank Alastair Smith, Hazel Knight, Judith Evans, and Geoff Foster for their work in the laboratory, Derek Brown and Dave Gally for input on phage dynamics, and Darren Shaw and Paul Bessel for help with the graphics.

This study was funded by the Wellcome Trust International Partnership Research Award in Veterinary Epidemiology (IPRAVE) project. SAC receives financial support from the Scottish Executive Environment and Rural Affairs Department (SEERAD). I.J.M. acknowledges the support of SEERAD project BSS/028/99. L.M. is grateful to the Wellcome Trust for a Mathematical Biology Research Training Fellowship.

REFERENCES

1. **Armstrong, G. L., J. Hollingsworth, and J. G. Morris.** 1996. Emerging food pathogens: *Escherichia coli* O157:H7 as a model entry of a new pathogen into the food supply of the developed world. Epidemiol. Rev. **18:**29–51.
2. **Blanco, J., M. Blanco, J. E. Blanco, A. Mora, M. P. Alonson, E. A. Gonzalez, and M. I. Bernardez.** 2001. Epidemiology of verocytotoxigenic *Escherichia coli* (VTEC) in ruminants, p. 113–148. *In* G. Duffy, P. Garvey, D. A. McDowell, and C. T. Trumbull (ed.), Verocytotoxigenic *E. coli*. Food and Nutrition Press, Inc., Trumbull, CT.
3. **Centers for Disease Control and Prevention.** 2005. Preliminary FoodNet data on the incidence of infection with pathogens transmitted commonly through food—10 sites, United States 2004. Morb. Mortal. Wkly. Rep. **54:** 352–356.
4. **Conedera, G., P. A. Chapman, S. Marangon, E. Tisato, P. Dalvit, and A. Zuin.** 2001. A field survey of *Escherichia coli* O157 ecology on a cattle farm in Italy. Int. J. Food Microbiol. **66:**85–93.
5. **Du, J.** 2002. Combined algorithms for constrained estimation of finite mixture distributions with grouped and conditional data. M.S. thesis. McMaster University, Ontario, Canada.
6. **Dziva, F., P. M. van Diemen, M. P. Stevens, A. J. Smith, and T. S. Wallis.** 2004. Identification of *Escherichia coli* O157:H7 genes influencing colonization of the bovine gastrointestinal tract using signature-tagged mutagenesis. Microbiology **150:**3631–3645.
7. **Fegan, N., P. Vanderlinde, G. Higgs, and P. Desmarchelier.** 2004. The prevalence and concentration of *Escherichia coli* O157 in faeces of cattle from different production systems at slaughter. J. Appl. Microbiol. **97:**362–370.
8. **Fitzpatrick, M. M., V. Shah, R. S. Trompeter, M. J. Dillon, and T. M. Barratt.** 1991. Long-term renal outcome of childhood haemolytic uremic syndrome. Br. Med. J. **303:**489–492.
9. **Galland, J. C., D. R. Hyatt, S. S. Crupper, and D. W. Acheson.** 2001. Prevalence, antibiotic susceptibility, and diversity of *Escherichia coli* isolates from a longitudinal study of beef cattle feedlots. Appl. Environ. Microbiol. **67:**1619–1627.
10. **Galvani, A. P., and R. M. May.** 2005. Epidemiology: dimensions of superspreading. Nature **438:**293–295.
11. **Garber, L. P., S. J. Wells, D. D. Hancock, M. P. Doyle, J. Tuttle, J. A. Shere, and T. Zhao.** 1995. Risk factors for faecal shedding of *Escherichia coli* O157:H7 in dairy calves. J. Am. Vet. Med. Assoc. **207:**46–49.
12. **Gilbert, M., A. Mitchell, D. Bourne, J. Mawdsley, R. Clifton-Hadley, and W. Wint.** 2005. Cattle movements and bovine tuberculosis in Great Britain. Nature **435:**491–496.
13. **Griffin, P. M., and R. V. Tauxe.** 1991. The epidemiology of infections caused by *Escherichia coli* O157:H7, other enterohemorrhagic *E. coli* and the associated haemolytic uremic syndrome. Epidemiol. Rev. **30:**60–98.
14. **Gunn, G. J., I. J. McKendrick, H. E. Ternent, F. Thomson-Carter, G. Foster, and B. A. Synge.** An investigation of factors associated with the prevalence of verocytotoxin producing *Escherichia coli* O157 shedding Scottish beef cattle. Vet. J., in press.
15. **Halliday, J. E. B., M. E. Chase-Topping, M. C. Pearce, I. J. Mckendrick, L. Allison, D. Fenlon, C. Low, D. J. Mellor, G. J. Gunn, and M. E. J. Woolhouse.** 2006. Herd-level risk factors associated with the presence of phage type 21/28 *E. coli* O157 on Scottish farms. BMC Microbiol. **6:**99.
16. **Hancock, D. D., T. E. Besser, M. L. Kinsel, P. I. Tarr, D. H. Rice, and M. G. Paros.** 1994. The prevalence of *Escherichia coli* O157:H7 in dairy and beef cattle in Washington State. Epidemiol. Infect. **113:**199–207.
17. **Hancock, D. D., T. E. Besser, D. H. Rice, D. E. Herriott, and P. I. Tarr.** 1997. A longitudinal study of *Escherichia coli* O157 in fourteen cattle herds. Epidemiol. Infect. **118:**193–195.
18. **Heuvelink, A. E., F. L. A. M. V. D. Biggelarr, E. de Boer, R. G. Herbes, W. J. G. Melchers, J. H. J. Huis In't Veld, and L. A. H. Monnens.** 1998. Isolation and characterization of verocytotoxin-producing *Escherichia coli* O157 strains from Dutch cattle and sheep. J. Clin. Microbiol. **36:**878–882.
19. **Hosmer, D. W., and S. Lemeshow.** 2000. Applied logistic regression, 2nd ed. Wiley, New York, NY.
20. **Innocent, G. T., D. J. Mellor, S. A. McEwen, W. J. Reilly, J. Smallwood, M. E. Locking, D. J. Shaw, P. Michel, D. J. Taylor, W. B. Steele, G. J. Gunn, H. E. Ternent, M. E. J. Woolhouse, and S. W. J. Reid.** 2005. Spatial and temporal epidemiology of sporadic human cases of *Escherichia coli* O157 in Scotland 1996–1999. Epidemiol. Infect. **153:**1033–1041.
21. **Karmali, M. A.** 1989. Infection by verocytotoxin-producing *Escherichia coli* O157. Clin. Microbiol. Rev. **2:**15–38.
22. **Karmali, M. A.** 2004. Infection by shiga toxin-producing *Escherichia coli*. An overview. Mol. Biotechnol. **26:**117–122.
23. **Khakhria, R., D. Duck, and H. Lior.** 1990. Extended phage-typing scheme for *Escherichia coli* O157:H7. Epidemiol. Infect. **105:**511–520.
24. **Kistemann, T., S. Zimmer, I. Vagsholm, and Y. Andersson.** 2004. GIS-supported investigation of human EHEC and cattle VTEC O157 infections in Sweden: geographical distribution, spatial variation and possible risk factors. Epidemiol. Infect. **132:**495–505.
25. **Lipsitch, M., T. Cohen, B. Cooper, J. M. Robins, S. Ma, L. James, G. Gopalakrishna, S. K. Chew, C. C. Tan, M. H. Samore, D. Fisman, and M. Murray.** 2003. Transmission dynamics and control of severe acute respiratory syndrome. Science **300:**1966–1970.
26. **Lloyd-Smith, J. O., S. J. Schreiber, P. E. Kopp, and W. M. Getz.** 2005. Superspreading and the effect of individual variation on disease emergence. Nature **438:**355–359.
27. **Locking, M., L. Browning, A. Smith-Palmer, and S. Brownlie.** 2005. Gastrointestinal and foodborne infections. Health Prot. Scotland Wkly. Rep. **39:** 2–4.
28. **Locking, M., L. Browning, A. Smith-Palmer, and S. Brownlie.** 2006. Gastrointestinal and foodborne infections. Health Prot. Scotland Wkly. Rep. **40:** 2–4.
29. **Locking, M. E., S. J. O'Brien, W. J. Reilly, D. M. Campbell, L. M. Browning, E. M. Wright, J. E. Coia, and J. E. Ramsay.** 2001. Risk factors for sporadic cases of *Escherichia coli* O157 infection: the importance of contact with animal excreta. Epidemiol. Infect. **127:**215–220.
30. **Low, J. C., I. J. McKendrick, C. McKechnie, D. Fenlon, S. W. Naylor, C. Currie, D. G. Smith, L. Allison, and D. L. Gally.** 2005. Rectal carriage of enterohemorrhagic *Escherichia coli* O157 in slaughtered cattle. Appl. Environ. Microbiol. **71:**93–97.
31. **Lynn, R. M., S. J. O'Brien, C. M. Taylor, G. K. Adak, H. Chart, T. Cheasty, J. E. Coia, I. A. Gillespie, M. E. Locking, W. J. Reilly, H. R. Smith, A. Waters, and G. A. Willshaw.** 2005. Childhood haemolytic uremic syndrome, United Kingdom and Ireland. Emerg. Infect. Dis. **11:**590–596.
32. **Matthews, L., J. C. Low, D. L. Gally, M. C. Pearce, D. J. Mellor, J. A. P. Heesterbeek, M. Chase-Topping, S. W. Naylor, D. J. Shaw, S. W. J. Reid, G. J. Gunn, and M. E. J. Woolhouse.** 2006. Heterogeneous shedding of *Escherichia coli* O157 in cattle and its implications for control. Proc. Natl. Acad. Sci. USA **103:**547–552.
33. **Matthews, L., I. J. McKendrick, H. Ternent, G. J. Gunn, B. Synge, and M. E. J. Woolhouse.** 2006. Super-shedding cattle and the transmission dynamics of *Escherichia coli* O157. Epidemiol. Infect. **134:**131–142.
34. **Meng, J. S., S. Zhao, M. P. Doyle, S. E. Mitchell, and S. Kresovich.** 1997. A multiplex PCR for identifying Shiga-like toxin-producing *Escherichia coli* O157:H7. Lett. Appl. Microbiol. **24:**172–176.
35. **Meyer-Broseta, S., S. N. Bastian, P. D. Arne, O. Cerf, and M. Sanaa.** 2001. Review of epidemiological surveys on the prevalence of contamination of

healthy cattle with *Escherichia coli* O157:H7. Int. J. Hyg. Environ. Health **203:**347–361.

36. **Michel, P., J. B. Wilson, S. W. Martin, R. C. Clarke, S. A. McEwan, and C. L. Gyles.** 1999. Temporal and geographic distributions of reported cases of *Escherichia coli* O157:H7 infection in Ontario. Epidemiol. Infect. **122:**193–200.
37. **Naylor, S. W., J. C. Low, T. E. Besser, A. Mahajam, G. J. Gunn, M. C. Pearce, I. J. McKendrick, D. G. E. Smith, and D. L. Gally.** 2003. Lymphoid follicle-dense mucosa at the terminal rectum is the principal site of colonisation of enterohemorrhagic *Escherichia coli* O157:H7 in the bovine host. Infect. Immun. **71:**1505–1512.
38. **Naylor, S. W., A. J. Roe, P. Nart, K. Spears, D. G. E Smith, J. C. Low, and D. L. Gally.** 2005. *E. coli* O157:H7 forms attaching and effacing lesions at the terminal rectum of cattle and colonization requires *LEE4* operon. Microbiology **151:**2773–2781.
39. **Nielsen, E. M., C. Tegtmeier, H. J. Andersen, C. Gronbaek, and J. S. Andersen.** 2002. Influence of age, sex and herd characteristics on the occurrence of verocytotoxin-producing *Escherichia coli* O157 in Danish farms. Vet. Microbiol. **88:**245–257.
40. **O'Brien, S. J., G. K. Adak, and C. Gilham.** 2001. Contact with farming environment as a major risk factor for shiga toxin (verocytotoxin)-producing *Escherichia coli* O157 infection in humans. Emerg. Infect. Dis. **7:**1049–1051.
41. **Ogden, I. D., M. MacRae, and N. J. C. Strachan.** 2004. Is the prevalence and shedding concentrations of *E. coli* O157 in beef cattle in Scotland seasonal? FEMS Microbiol. Lett. **233:**297–300.
42. **Omisakin, F., M. MacRae, I. D. Ogden, and N. J. C. Strachan.** 2003. Concentration and prevalence of *Escherichia coli* O157 in cattle feces at slaughter. Appl. Environ. Microbiol. **69:**2444–2447.
43. **Paiba, G. A., J. W. Wilesmith, S. J. Evans, S. J. S. Pascoe, R. P. Smith, S. A. Kidd, J. B. M. Ryan, I. M. McLaren, S. A. Chappell, G. A. Willshaw, T. Cheasty, N. P. French, T. W. H. Jones, H. F. Buchanan, D. J. Challoner, A. D. Colloff, M. P. Cranwell, R. G. Daniel, I. H. Davies, J. P. Duff, R. A. T. Hogg, F. D. Kirby, M. F. Millar, R. J. Monies, M. J. Nicholls, and J. H. Payne.** 2003. Prevalence of faecal excretion of verocytotoxogenic *Escherichia coli* O157 in cattle in England and Wales. Vet. Rec. **153:**347–353.
44. **Pearce, M. C., D. Fenlon, J. C. Low, A. W. Smith, H. I. Knight, J. Evans, G. Foster, B. A. Synge, and G. J. Gunn.** 2004. Distribution of *Escherichia coli* O157 in bovine fecal pats and its impact on estimates of the prevalence of fecal shedding. Appl. Environ. Microbiol. **70:**5737–5743.
45. **Robinson, S. E., E. J. Wright, C. A. Hart, M. Bennett, and N. P. French.** 2004. Intermittent and persistent shedding of *Escherichia coli* O157 in cohorts of naturally infected calves. J. Appl. Microbiol. **97:**1045–1053.
46. **Roe, A. J., S. W. Naylor, K. J. Spears, H. M. Yull, T. A. Dransfield, M. Oxford, I. Mckendrick, M. Porter, M. Woodward, D. G. E. Smith, and D. L. Gally.** 2004. Co-ordinate single cell expression of LEE4- and LEE5-encoded proteins of *Escherichia coli* O157:H7. Mol. Microbiol. **54:**337–352.
47. **RR Development Core Team.** 2005. R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. URL http://www.R-project.org.
48. **Sargeant, J. M., M. W. Sanderson, R. A. Smith, and D. D. Griffin.** 2003. *Escherichia coli* O157 in feedlot cattle faeces and water in four major feedercattle states in the USA. Prev. Vet. Med. **61:**127–135.
49. **Schouten, J. M., M. Bouwknegt, A. W. van de Giessen, K. Frankena, M. C. M. De Jong, and E. A. M. Graat.** 2004. Risk factor analysis of verocytotoxin producing *Escherichia coli* O157 on Dutch dairy farms. Prev. Vet. Med. **64:**49–61.
50. **Shere, J. A., K. J. Bartlett, and C. W. Kaspar.** 1998. Longitudinal study of *Escherichia coli* O157:H7 dissemination on four dairy farms in Wisconsin. Appl. Environ. Microbiol. **64:**1390–1399.
51. **Strachan, N. J., G. M. Dunn, M. E. Locking, T. M. S. Reid, and I. D. Ogden.** 2006. *Escherichia coli* O157: burger bug or environmental pathogen. J. Food Microbiol. **112:**129–137.
52. **Willshaw, G. A., S. M. Scotland, H. R. Smith, T. Cheasty, A. Thomas, and B. Rowe.** 1994. Hybridization of strains of *Escherichia coli* O157 with probes derived from the *eae* gene of enteropathogenic *Escherichia coli* and the *eae* homolog from a verocytotoxin-producing strain of *Escherichia coli* O157. J. Clin. Microbiol. **32:**897–902.
53. **Willshaw, G. A., J. Evans, T. Cheasty, A. Cummins, and G. C. Pritchard.** 2003. Verocytotoxin-producing *Escherichia coli* infection and private farm visits. Vet. Rec. **152:**365–366.
54. **Wilson, J. B., S. A. McEwen, R. C. Clarke, K. E. Leslie, D. Waltner-Toews, and C. L. Gyles.** 1993. Risk factors for bovine infection with verocytotoxigenic *Escherichia coli* in Ontario. Can. J. Prev. Vet. Med. **16:**159–170.
55. **Woolhouse, M. E. J., C. Dye, J.-F. Etard, T. Smith, J. D. Charlwood, G. P. Garnett, P. Hagan, J. L. K. Hii, P. D. Ndhlovu, R. J. Quinnell, C. H. Watts, S. K. Chandiwana, and R. M. Anderson.** 1997. Heterogeneities in the transmission of infectious agents: implications for the design of control programs. Proc. Natl. Acad. Sci. USA **94:**338–342.
56. **Zhao, T., M. P. Doyle, J. Shere, and L. Garber.** 1995. Prevalence of enterohemorrhagic *Escherichia coli* O157:H7 in a survey of dairy herds. Appl. Environ. Microbiol. **61:**1290–1293.