Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Preventive Veterinary Medicine 71 (2005) 45–56
PREVENTIVE VETERINARY MEDICINE
www.elsevier.com/locate/prevetmed

Spatial distribution of *Escherichia coli* O157-positive farms in Scotland

Nils Toft [a,*], Giles T. Innocent [b], Iain J. McKendrick [d], Helen E. Ternent [b], Dominic J. Mellor [b], George J. Gunn [e], Barti Synge [e], Stuart W.J. Reid [b,c]

[a] *Department of Large Animal Sciences, The Royal Veterinary and Agricultural University, Grønnegårdsvej 8, DK-1870 Frederiksberg C, Denmark*
[b] *Comparative Epidemiology & Informatics, Institute of Comparative Medicine, Faculty of Veterinary Medicine, University of Glasgow, Bearsden Road, Glasgow G61 1QH, UK*
[c] *Department of Statistics and Modeling Science, University of Strathclyde, Richmond Street, Glasgow G1 1XW, UK*
[d] *Biomathematics & Statistics Scotland, The King's Buildings, Edinburgh EH9 3JZ, UK*
[e] *SAC Veterinary Science Division, Drummondhill, Stratherick Road, Inverness IV2 4JZ, UK*

Received 11 July 2004; received in revised form 18 May 2005; accepted 23 May 2005

Abstract

Using a sample of 949 Scottish farms with finishing cattle, the spatial distribution of *Escherichia coli* O157-positive farms was investigated using disease mapping models. The overall prevalence of *E. coli* O157-positive farms was estimated as 22%. The regions used in this study were the 16 postcode areas of Scotland. For each region, the posterior relative risk (RR) was estimated as a model-based alternative to the saturated standardized morbidity ratio (SMR), i.e., the ratio between observed and expected cases in a region. Three Bayesian hierarchical models with generalized linear modeling of the area-specific risks were used to estimate the posterior relative risk of *E. coli* O157-positive farms in the postcode areas: a random-effects model incorporating only spatially uncorrelated heterogeneity; a model incorporating both spatially correlated and uncorrelated heterogeneity; and a pseudo-mixture model with unstructured correlation and a weighted mix of two variance components representing the spatial correlation and a jump structure. None of the models identified any areas with

* Corresponding author. Tel.: +45 35283024; fax: +45 35283055.
E-mail address: nt@dina.kvl.dk (N. Toft).

0167-5877/$ – see front matter
doi:10.1016/j.prevetmed.2005.05.010

a significant increase or decrease in risk. The deviance information criteria slightly favored the simplest model (RR range: 0.92–1.09). However, this model appeared to smooth out more of the variation in the RR compared to the pseudo-mixture model, which gave a more informative pattern of the posterior relative risks (range: 0.81–1.22).

Keywords: Standardized morbidity ratio (SMR); Relative risk; Disease mapping; Bayesian hierarchical models

1. Introduction

Escherichia coli belonging to serogroup O157 can be an important human diarrhoeal pathogen. Cattle seem mostly unaffected by *E. coli* O157, but undercooked or raw meat is a source for food-borne infection in humans. Visits to infected farms, contact with animal excreta and recreational use of animal pasture are risk factors contributing to the sporadic cases of *E. coli* O157 in humans (Belongia et al., 2003; Locking et al., 2001; O'Brien et al., 2001; Strachan et al., 2002).

In analyses of the spatial distribution of human cases, incidences of human cases were associated with higher values of various indicators of livestock intensity (Innocent et al., 2005; Kistemann et al., 2004; Michel et al., 1999; Valcour et al., 2002). However, no studies have investigated the spatial distribution of *E. coli* O157-positive farms on a large scale.

To help control the potential transmission risk of *E. coli* O157 from cattle to humans, mapping of *E. coli* O157-positive farms and the assessment of the relative prevalence in smaller areas within the area under study, e.g., a country or a region, would be useful. If geographical differences in the prevalence of *E. coli* O157-positive farms exist, then concentrating on the high-risk areas could optimize the use of the scarce resources of a control program. More fundamentally, identification of spatial structures in the incidence of an infection, such as *E. coli* O157, would lead to an increased understanding of the disease/infection in terms of geographically associated risk factors.

A study of the farm-level prevalence of *E. coli* O157 in Scottish fattening cattle found 22% of all herds *E. coli* O157-positive (Synge et al., 2001; Ternent, 2002). In that study, Scotland was divided into six areas based on the Scottish Agricultural College's Animal Health Divisions to investigate potential differences in the prevalence of *E. coli* O157-positive farms, but no significant effect was found. However, that study did not allow for possible spatial variation within regions. Furthermore, dividing Scotland into only six regions might result in too crude a representation of potential spatial structures relevant to the prevalence *E. coli* O157-positive farms.

Our objective was to characterize the spatial distribution of *E. coli* O1 57-positive farms in Scotland using the relative risk of *E. coli* O157-positive farms in smaller areas formed from postcode-information. As simple models might smooth over large discontinuities in the risk surface (Lawson and Clark, 2002), we used three different models: a simple random-effects model with an unstructured effect; a model where spatially structured effects were included; and a model which also allowed discrete jumps in the underlying spatial structure.

2. Materials

We used data from a study that was designed to estimate the prevalence of *E. coli* O157-positive fattening cattle-farms in Scotland (Ternent, 2002). The study was also designed to investigate a variety of potential risk factors, including the effect of Animal Health Division (AHD) as a proxy for potential geographical differences. Thus, sampling was stratified by AHD to ensure that the samples were representative of the distribution of farms across individual AHDs. We assume that the samples were representative of the spatial distributions at lower levels of spatial aggregation as well, this assumption being justified by the randomness of the sample at these levels. Farms were selected at random using computer-generated lists from a database of all Scottish cattle fattening farms.

On each farm, a number of faecal pats were sampled. The number of pats sampled on each farm was determined by a sampling scheme designed to achieve an 80% probability of identifying the sampled group as positive, if at least one infected animal was shedding. The sampling scheme was based on the size of the group of finishing cattle and an assumption about the likely mean within-group prevalence. The latter was estimated as being approximately 10% using data from an unpublished pilot-study conducted by two of the authors.

Faecal samples were examined for *E. coli* O157 strains using immunomagnetic separation (IMS) as described by Chapman et al. (1994), but using buffered peptone water (BPW) without added antibiotics during enrichment to increase the analytical sensitivity (Foster et al., 2003).

The sampled farms were grouped into 16 postcode areas of Scotland (Fig. 1(a) and (b)). The sampling was carried out over 2 years, but any potential temporal structure is ignored in this study due to the relatively small sample sizes in each postcode area.

3. Methods

The prevalence (p_i) of *E. coli* O157-positive farms in the ith area was estimated as $p_i = y_i/n_i$, where y_i is the number of positive farms in the area and n_i the total number of farms sampled. A farm was defined as *E. coli* O157-positive if at least one pat sample from the farm was positive.

The standardized morbidity ratio (SMR_i) for the ith area is defined as the ratio between the observed and expected cases. Thus, the SMR can be defined as the ratio between the area-specific prevalence (p_i) and the overall prevalence, i.e. $p = y/n$, where $y = \sum y_i$ and $n = \sum n_i$. The SMR is often used in disease mapping, but has many disadvantages (Lawson et al., 2003). The SMR is a saturated estimate of the relative risk, where relatively small changes in the expected value can yield large changes in estimates. When zero cases are observed, the SMR will be zero, regardless of the expected number of cases. Furthermore, the variance of the SMR is proportional to the expected number of positives.

To address the problems associated with SMRs, we adopted model-based relative risk estimation methods to smooth the SMR estimates. As an initial model, we used a

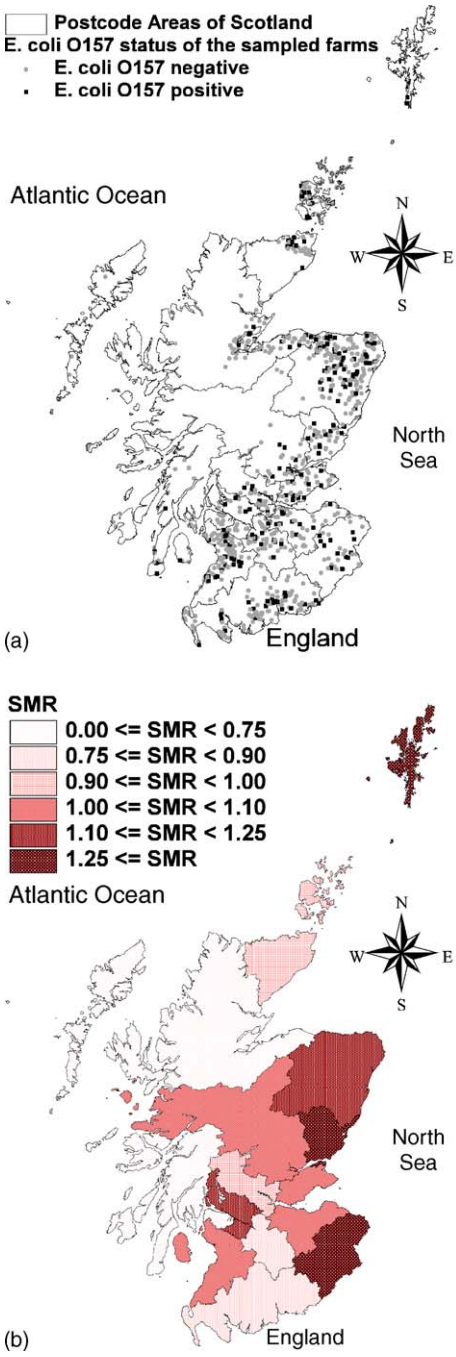

Fig. 1. (a) The spatial distribution of the 949 sampled farms in Scotland; (b) the standardized morbidity ratio (SMR) of the prevalence of *E. coli* O157-positive farms.

generalized linear model with a logit link-function with random-effects:

$$y_i \sim \text{Bin}(p_i, n_i), \qquad \text{logit}(p_i) = \alpha + v_i, \qquad v_i \sim N(0, \sigma_v^2) \tag{1}$$

where y_i, n_i and p_i were the number of *E. coli* O157-positive farms, total number of sampled farms and estimated prevalence in the ith postcode area, and v_i were the random-effects associated with postcode area, modeling the (uncorrelated) heterogeneity between postcode areas. The parameter α, modeling the common intercept defining the mean global-prevalence, was given the improper prior, dflat, in WinBUGS, which can be interpreted as a uniform distribution on the real line (Spiegelhalter et al., 2003). This improper distribution on α was needed in the two models specified below and used in this model to facilitate comparisons between models. The random-effects parameter σ_v was given a non-informative uniform prior distribution on the interval [0.01, 10]. This choice of prior allowed σ_v to vary uniformly in the relevant interval, while prohibiting extreme values which caused problems with the convergence of the MCMC chain.

The relative risk RR_i of the ith area was then estimated using $\text{RR}_i = p_i/p$, where p is the overall prevalence. The RR and SMR represents the same estimator, applied to smoothed and unsmoothed data, respectively, but we will refer to the model-based statistics as RR and the saturated statistics as SMR to distinguish between the results.

Although the above model smooths the estimates of relative risk, it does not explicitly incorporate any spatial structure. The Besag, York and Mollié (BYM) model (Besag et al. (1991) cited from Lawson et al. (2003)) extends the random-effects model by decomposing the area-specific random-effects into a spatially structured effect (clustering or correlated heterogeneity) and spatially unstructured variability. Formally, the model is specified by:

$$\begin{aligned} &y_i \sim \text{Bin}(p_i, n_i), \qquad \text{logit}(p_i) = \alpha + v_i + u_i, \qquad v_i \sim N(0, \sigma_v^2), \\ &(u_i | u_j, i \neq j, \sigma_u^2) \sim N(\bar{u}_i, \sigma_i^2) \end{aligned} \tag{2}$$

where

$$\bar{u}_i = \frac{1}{\sum_j \psi_{ij}} \sum_j u_j \psi_{ij}, \qquad \sigma_i^2 = \frac{\alpha_u^2}{\sum_j \psi_{ij}}, \qquad \psi_{ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are adjacent,} \\ 0 & \text{otherwise.} \end{cases}$$

i.e., spatial correlation is explicitly modeled, so that the estimation of risk parameters in any area is conditional on the values estimated in neighboring areas. This model is unidentifiable and it is necessary to impose a constraint on, for example, the intercept α through the use of an improper uniform distribution as discussed above. The random-effects parameters σ_v and σ_u are given uniform priors on the interval [0.01, 10] as before.

Both the random-effects model and the BYM model can smooth out genuine and relevant discontinuities. Hence Lawson and Clark (2002) suggested an extension of the BYM model to allow discrete jumps in prevalence. Lawson and Clark referred to their model as a mixture model; however, that term is usually reserved for models which are defined as a mixture of the distributions of mutually exclusive sub-populations, where one of these represents the true distribution of the subject (e.g. whether a postcode area is similar or dissimilar to its neighbors) (Gelman et al., 1994). The model proposed in Lawson and Clark (2002) assumes that the spatial variation in each postcode area is a weighted mix

of the spatially structured effect of the BYM model and a jump-process, i.e., all areas contain both components but in varying proportions. We will refer to this approach as a pseudo-mixture model. Formally, the pseudo-mixture model is defined as follows:

$$y_i \sim \text{Bin}(p_i, n_i), \qquad \text{logit}(p_i) = \alpha + v_i + \rho_i u_i + (1 - \rho_i) w_i \tag{3}$$

where v_i and u_i are defined as for the BYM model, and w_i by the joint distribution $\pi(w_1, w_2, \ldots, w_m)$

$$\pi(w_1, w_2, \ldots, w_m) \propto \frac{1}{\sqrt{\lambda}} \exp\left[-\frac{1}{\lambda} \sum_{i,j} \psi_{ij} |w_i - w_j| \right]$$

where λ acts as a constrained term with a gamma prior distribution, and the weights ρ_i are given the prior distribution Beta(0.5, 0.5), i.e., a prior belief that the weights attached to the two underlying processes will be disparate rather than similar, but with no assumption of either process being dominant. The jump-component of the pseudo-mixture model is defined in terms of the total absolute difference between the jump-model variance-components in neighboring areas. While the choice of using differences between risks in neighboring areas seems obvious, the choice of functional relationship, i.e., the use of total absolute difference, is less obvious and essentially chosen for convenience of implementation. Again, α must be assigned an improper prior to allow identifiability of the spatial heterogeneity element of the model, while the prior distributions of the random-effects σ_v, σ_u and σ_w are assumed uniform on the interval [0.01, 10].

The relative risk using each of the three models were estimated using WinBUGS (Spiegelhalter et al., 2003). For each model an initial burn-in of 5000 iterations were followed by sampling a chain of 50,000 samples which were thinned down to every 10th sample to reduce auto-correlation. Convergence was assessed by time-series plots of the thinned chains.

The three models were compared using their Bayesian residuals (r_i, defined as $r_i = y_i/n_i - \hat{p}_i$, where $\hat{p}_i$ is the estimated prevalence) and the deviance information criterion (DIC) (Spiegelhalter et al., 2002).

4. Results

In total, 952 farms were sampled. However, due to missing or incorrect geographical information, three farms were excluded from further analysis in the present study. The spatial distribution of the remaining 949 farms is shown in Fig. 1(a) using the geographical coordinates recorded using Ordnance Survey maps at the farm visit. The farms represented a sample of approximately 10% of Scottish farms with finishing beef cattle during the sampling period. Of these 949 farms, 213 were identified as positive, an overall farm-level prevalence of 22%. The large area without samples in the northwest of Scotland is an area where the primary livestock production is sheep and there are very few farms fattening cattle.

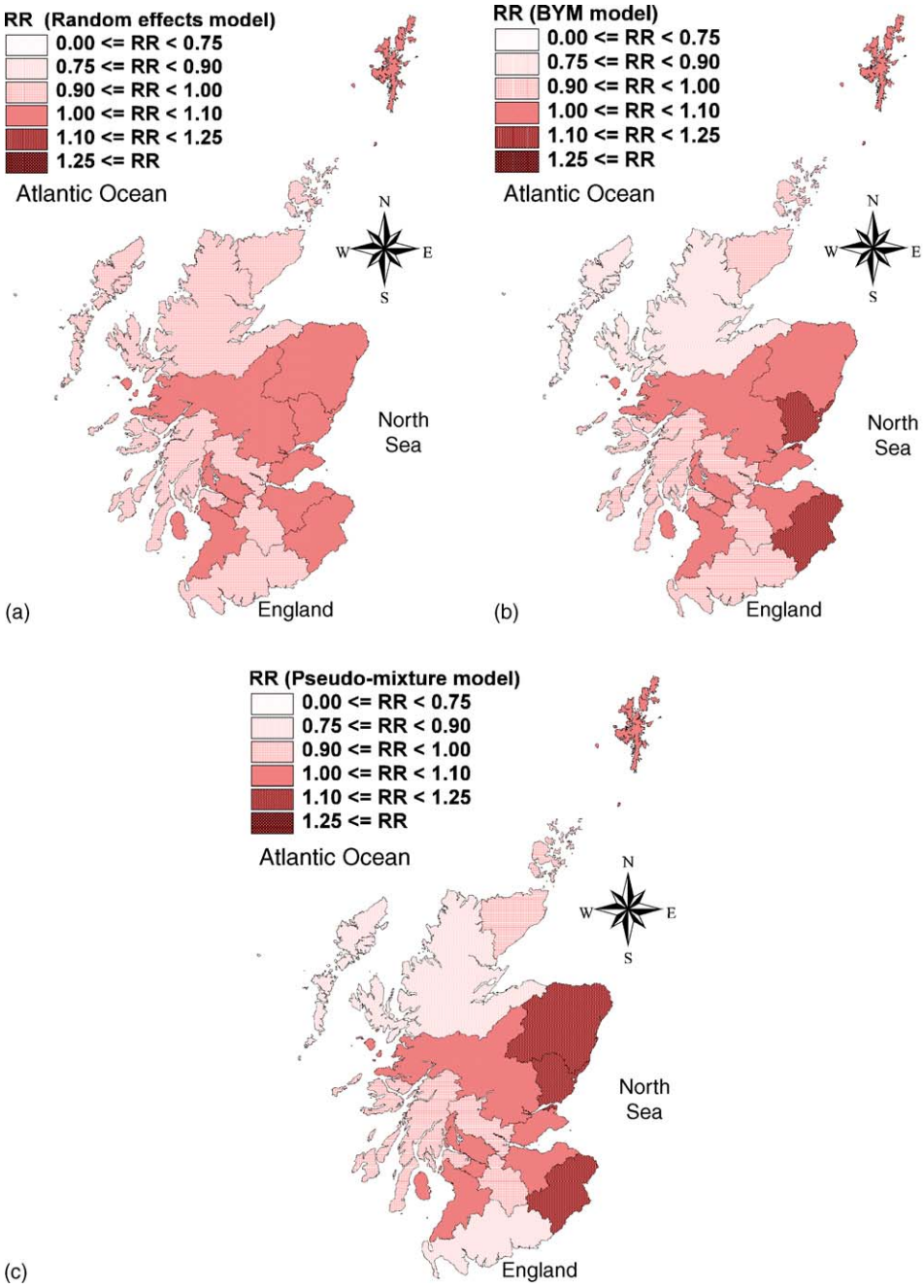

Fig. 2. The posterior expected relative risk (RR) for each of the three models: (a) the random-effects model; (b) the BYM model; (c) the pseudo-mixture model.

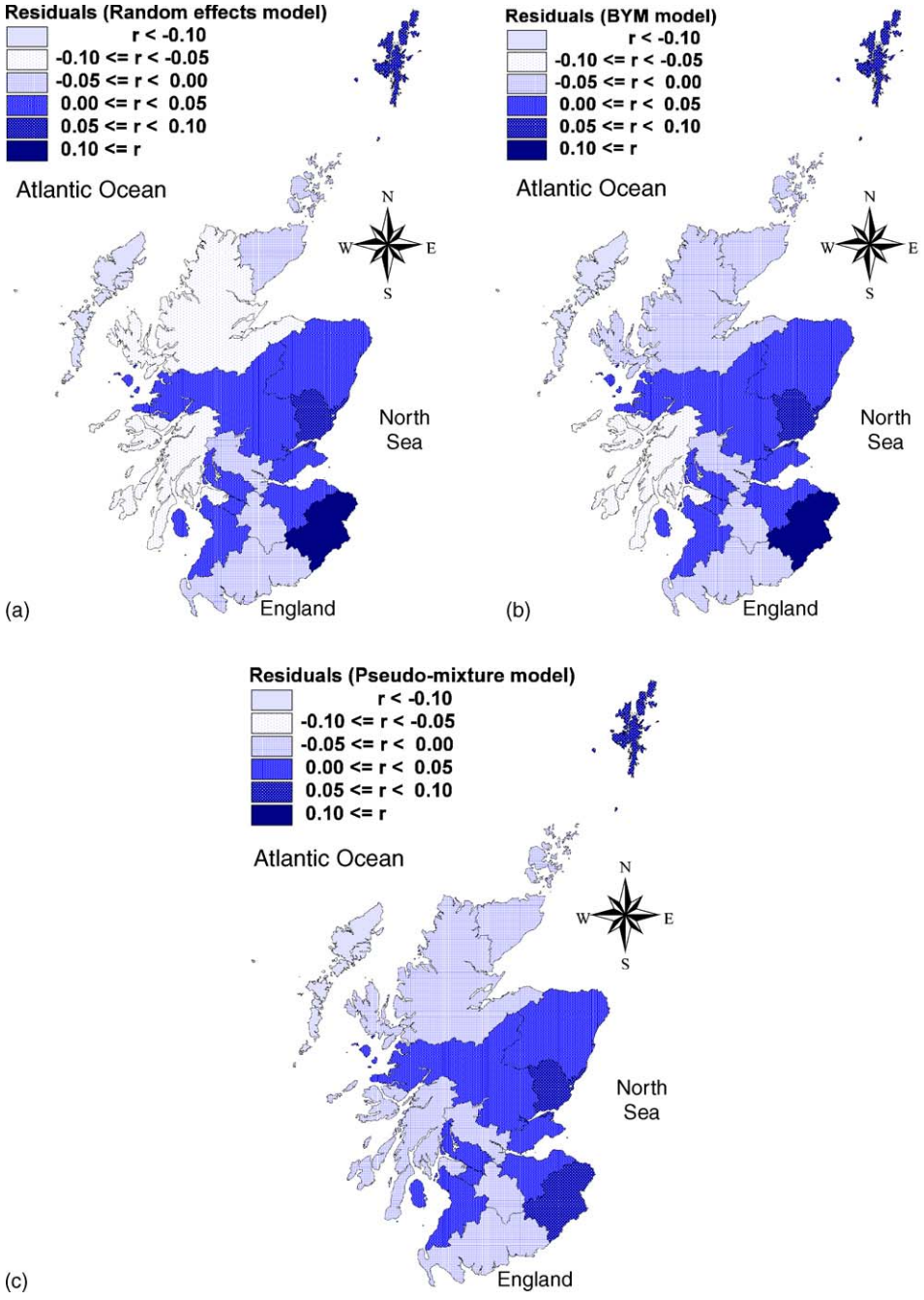

Fig. 3. The residuals for each of the three models: (a) the random-effects model; (b) the BYM model; (c) the pseudo-mixture model.

Table 1
The deviance information criteria (DIC), the posterior estimated mean and 95% credibility posterior interval, given as the estimated 2.5 and 97.5% percentiles, of the population mean prevalence (p), the intercept α and variance components (σ_v, σ_u and σ_w) for the three RR models

	Random-effects model			BYM model			Pseudo-mixture model		
	2.5%	Mean	97.5%	2.5%	Mean	97.5%	2.5%	Mean	97.5%
DIC		79.91			81.29			80.93	
p	0.19	0.22	0.26	0.19	0.22	0.26	0.18	0.22	0.26
α	−1.45	−1.25	−1.07	−1.46	−1.26	−1.07	−1.50	−1.27	−1.05
σ_v	0.02	0.16	0.42	0.02	0.16	0.44	0.02	0.17	0.47
σ_u	–	–	–	0.02	0.27	0.80	0.03	0.53	1.77
σ_w	–	–	–	–	–	–	0.07	0.58	1.35

The SMR of *E. coli* O157-positive farms by postcode area is given in Fig. 1(b). The SMR (range: 0.00–1.62) suggested that there was some spatial variation in the distribution of *E. coli* O157-positive farms.

The posterior expected relative risk for farms being *E. coli* O157-positive using the simple random-effects model is shown in Fig. 2(a). The random-effects model (Fig. 2(a)), range of the posterior expected RR: 0.92–1.09, removed most of the variation seen in the SMR (Fig. 1(b)). The posterior probabilities of the RRs being different from one ($\Pr(RR \neq 1)$) were below 71%, i.e. the RRs for the 16 postcode areas could not be considered different from one at the usual 5% significance level.

As could be expected, the posterior expected relative risk (Fig. 2(b)) for the BYM model showed less smoothing compared to Fig. 2(a). The posterior expected RR estimates ranged from 0.87 to 1.16, with $\Pr(RR \neq 1)$ less than 82%, i.e. no RR were significantly different from one at the 5% significance level.

The pseudo-mixture model with discrete jumps as well as a spatial structure gave the least-smoothed pattern of the three models (Fig. 2(c)). By allowing the discrete jumps in the model, the posterior expected RR estimates ranged from 0.81 to 1.22, with the posterior $\Pr(RR \neq 1)$ less than 84%, i.e. no RR were different from one at the 5% significance level.

The posterior estimates of the population mean-prevalence $p = \exp(\alpha)/(1 + \exp(\alpha))$, α and the variance components (σ_v, σ_u, σ_w) for each of the three models are given in Table 1 as the posterior estimated mean with the estimated 2.5 and 97.5% percentiles. The deviance information criteria (DIC) is also given for each model; the minimum DIC estimates the model that will make the best short-term predictions.

Fig. 3 shows the Bayesian residuals for each of the three models; as expected, the overall range of the residuals decreases as the complexity of the models increases.

5. Discussion and conclusion

The posterior estimated means and 95% credible posterior intervals for the intercept α and the unstructured variance component σ_v are similar across the three models (Table 1), which seems to indicate that the models behave reasonably well.

Lawson et al. (2003) suggests the use of the Bayesian residuals as a means of assessing the local goodness-of-fit. Specifically, the ranges of the residuals are used to indicate which model has the better fit, while the patterns of the residuals identify potential areas where the models have problems estimating the true relative risk. The residual plots (Fig. 3) suggest that there might be something gained by applying the more complicated models, because the map for the pseudo-mixture model (Fig. 3(c)) shows residuals with a smaller range than the random-effects model (Fig. 3(a)). However, all three plots display residuals with a reasonable range indicating a reasonable fit. The plots all show that there are relatively larger problems estimating the true relative risk in the north west and the south east of Scotland, but that the problems decrease as the complexity of the models increase.

The DIC in Table 1 favors the simplest model over the pseudo-mixture model and the BYM model, with the BYM as the least favorable model. However, according to Spiegelhalter et al. (2003), if the difference in DIC is, say, less than five, and the models make very different inferences, then it could be misleading just to report the model with the lowest DIC. Hence, because any differences in DIC and residual maps are not very large there is no strong evidence against any of the models. A model fitting only the intercept α gives a DIC of 78.50, and could therefore be preferred to any of these three models, but, as the differences in DIC are all less than three, we chose to take the above advice and consider the nature of the inferences from the different models.

None of the three models identified any clusters with a significant reduction or increase in relative risk compared to the overall risk level, at the 5% significance level. The pattern in the pseudo-mixture model (Fig. 2(c)) did, however, show more variation than the other two models. Prior to the analysis carried out in this study, we analyzed data for clusters of *E. coli* O157-positive farms using the (x,y)-coordinates of the individual farms to calculate scan statistics using SaTScan v4.0 (http://www.satscan.org) (Kulldorf, 1997). This analysis did not find any significant clusters, which is consistent with the findings here. However, simple sample-size calculations based on a base prevalence of 0.22 and RR = 1.25 suggest that 760 farms should be sampled in each (of two) areas to find a such difference between the two areas with 95% confidence and 80% power (Houe et al., 2004). Hence, failing to identify any significant differences in this study is hardly surprising considering that only 949 farms in total were sampled.

A Swedish study (Kistemann et al., 2004) reports that positive cattle samples appeared to be concentrated in the southern and central part of Sweden. However, the study reported only the cases and did not adjust for the spatial distribution of farms in general, so it is impossible to make any inference from that study about the spatial distribution of positive farms.

Analyzing the data on the spatial scale of the 16 postcode areas does not give much scope for modeling the spatial structure. This scale was chosen to avoid the problems with convergence due to small sample sizes and missing information in areas that occur when using finer scales of postcode-information such as postcode-districts (of which there are 434 in Scotland). A more suitable spatial scale would probably be something intermediate between the two scales and an optimal choice will depend on the objective of the study. However, because no clusters were identified using individual farm records, it is unlikely that regions with significant deviations in risk will be identified at any level of aggregation, especially given the sample-size considerations outlined above. It is expected that using a

spatial scale with more regions would favor more-complex models, as spatial heterogeneity becomes more important when smoothing at a less aggregated scale; dividing an area into more regions can reveal a structure that is invisible at a higher level of aggregation.

To utilize the RR for inference and decision making we would suggest that the pseudo-mixture model is used. To discover potential sub-populations the model needs to be formulated in a way which allows such populations to be identified (McCulloch et al., 2002). Future analyses of the data presented in this study should involve ecological analyses of covariates at area level, e.g., indicators of livestock density, as well as farm level covariates, such as management related factors. In the present study the pseudo-mixture model did not identify areas with significantly higher or lower risk. Furthermore, based on the RR estimates, the calculated smoothed area-specific risks lies between 18 and 27%. This basically suggests that differences are not only non-significant, but also too small to advocate an area-specific intervention against *E. coli* O157 among Scottish cattle fattening farms.

An advantage of the models used here compared to models based on individual farm recordings lies in the need for data quality. Geocoding the positions of farms is tedious, expensive, time consuming and in some countries been resisted by farmers opposing the potential for surveillance arising from such registration. However, census data can often be obtained at some level of aggregation, thus making the use of RR models a more practical and realistic option.

Acknowledgements

The original sampling survey was funded by SERAD (now SEERAD). This analysis was funded as part of the Wellcome Trust International Partnership Research Awards in Veterinary Epidemiology. Nils Toft was supported by a grant from The Danish Research Agency. Iain McKendrick was supported by SEERAD project BSS/028/99.

The maps were 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.

References

Belongia, E.A., Chyou, P.-H., Greenlee, R.T., Perez-Perez, G., Bibb, W.F., DeVries, E.O., 2003. Diarrhea incidence and farm-related risk factors for *Escherichia coli* O157:H7 and *Campylobacter jejuni* antibodies among rural children. J. Infect. Dis. 187, 1460–1468.

Besag, J., York, J., Mollié, A., 1991. Bayesian image restoration with two applications in spatial statistics. Ann. Inst. Stat. Math. 43, 1–59.

Chapman, P.A., Wright, D.J., Siddons, C.A., 1994. A comparison of immunomagnetic separation and direct culture for the isolation of verototox-inproducing *Escherichia coli* O157 from bovine feces. J. Med. Microbiol. 40, 424–427.

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 preenrichment media prior to immunomagnetic separation for the isolation of *E. coli* O157 from bovine faeces. J. Appl. Microbiol. 95, 155–159.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 1994. Bayesian Data Analysis, 1st ed., Texts in Statistical Science Series. Chapman & Hall.

Houe, H., Ersbøll, A.K., Toft, N. (Eds.), 2004. Introduction to Veterinary Epidemiology. Biofolia.

Innocent, G.T., Mellor, D.J., McEwen, S.A., Reilly, W.J., Smallwood, J., Locking, M.E., Shaw, D.J., Michel, P., Taylor, D.J., Steele, W.B., Gunn, G.J., Ternent, H.E., Woolhouse, M.E.J., Reid, S.W.J., 2005. Spatial and temporal epidemiology of sporadic human cases of *Escherichia coli* O157 in Scotland, 1996–1999. Epidemiol. Infect., pp. 1–9.

Kistemann, T., Zimmer, S., Vågsholm, I., Andersson, Y., 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 (03), 495–505.

Kulldorf, M.A., 1997. A spatial scan statistic. Commun. Stat.: Theory Meth. 26, 1481–1496.

Lawson, A.B., Browne W.J, Vidal Rodeiro, C.L., 2003. Disease mapping with WinBUGS and MLwinN. Statistics in Practice, Wiley.

Lawson, A.B., Clark, A., 2002. Spatial mixture relative risk models applied to disease mapping. Stat. Med. 21, 359–370.

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

McCulloch, C.E., Lin, H., Slate, E.H., Turnbull, B.W., 2002. Discovering sub-population structure with latent class mixed models. Stat. Med. 21, 417–429.

Michel, P., Wilson, J.B., Martin, S.W., Clarke, R.C., McEwen, S.A., Gyles, C.L., 1999. Temporal and geographical distributions of reported cases of *Escherichia coli* O157:H7 infection in Ontario. Epidemiol. Infect. 122, 193–200.

O'Brien, S.J., Adak, G.K., Gilham, C., 2001. Contact with farming environment as a major risk factor for shiga toxin (vero cytotoxin)-producing *Escherichia coli* O157 infection in humans. Emerg. Infect. Dis. 7 (6), 1049–1051.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., van der Linde, A., 2002. Bayesian measures of model complexity and fit (with discussion). J. Roy. Stat. Soc., Ser. B: Stat. Methodol. 64 (4), 583–639.

Spiegelhalter, D.J., Thomas, A., Best, N.G., 2003. WinBUGS Version 1.4 User Manual. MRC Biostatistics Unit.

Strachan, N.J.C., Dunn, G.M., Ogden, I.D., 2002. Quantitative risk assessment of human infection from *Escherichia coli* O157 associated with recreational use of animal pasture. Int. J. Food Microbiol. 75, 39–51.

Synge, B.A., Gunn, G.J., Ternent, H.E., Hopkins, G.F., Thomson-Carter, F., Foster, G., Chase-Topping, M., McKendrick, I., 2001. Prevalence and factors affecting the shedding of verototoxin-producing *Escherichia coli* O157 in beef cattle in Scotland. In: Duffy, G., Garvey, P., Wasteson, Y., McDowell, D.A. (Eds.), Proceedings of a Meeting Hosted at Teagasc, The National Food Centre, at The Grand Hotel and Conference Centre, Malahide, Dublin, Ireland, 8–10 February 2001, pp. 98–103.

Ternent, H.E., 2002. An epidemiological study of *Escherichia coli* O157 strains in Scottish fattening cattle. Master Thesis. Faculty of Medicine, University of Aberdeen, University of Aberdeen, Academic Section, University Office, Regent Walk, Aberdeen AB24 3FX, UK.

Valcour, J.E., Michel, P., McEwene, S.A., Wilson, J.B., 2002. Associations between indicators of livestock farming intensity and incidence of human shiga toxin-producing *Escherichia coli* infection. Emerg. Infect. Dis. 8 (3), 252–257.