**Original article**

# A comparison of the active surveillance of scrapie in the European Union

Victor Javier DEL RIO VILAS[1]*, Dankmar BÖHNING[2], Ronny KUHNERT[3]

[1] Veterinary Laboratories Agency, New Haw, Addlestone, Surrey, KT15 3NB, UK
[2] Applied Statistics, School of Biological Sciences, The University of Reading, Harry Pitt Building, Earley Gate, Whiteknights road, Reading, RG6 6FN, UK
[3] Department of Epidemiology and Health, Robert Koch-Institute, Seestraβe 10, 13353 Berlin, Germany

**Abstract** – The abattoir and the fallen stock surveys constitute the active surveillance component aimed at improving the detection of scrapie across the European Union. Previous studies have suggested the occurrence of significant differences in the operation of the surveys across the EU. In the present study we assessed the standardisation of the surveys throughout time across the EU and identified clusters of countries with similar underlying characteristics allowing comparisons between them. In the absence of sufficient covariate information to explain the observed variability across countries, we modelled the unobserved heterogeneity by means of non-parametric distributions on the risk ratios of the fallen stock over the abattoir survey. More specifically, we used the profile likelihood method on 2003, 2004 and 2005 active surveillance data for 18 European countries on classical scrapie, and on 2004 and 2005 data for atypical scrapie separately. We extended our analyses to include the limited covariate information available, more specifically, the proportion of the adult sheep population sampled by the fallen stock survey every year. Our results show that the between-country heterogeneity dropped in 2004 and 2005 relative to that of 2003 for classical scrapie. As a consequence, the number of clusters in the last two years was also reduced indicating the gradual standardisation of the surveillance efforts across the EU. The crude analyses of the atypical data grouped all the countries in one cluster and showed non-significant gain in the detection of this type of scrapie by any of the two sources. The proportion of the population sampled by the fallen stock appeared significantly associated with our risk ratio for both types of scrapie, although in opposite directions: negative for classical and positive for atypical. The initial justification for the fallen stock, targeting a high-risk population to increase the likelihood of case finding, appears compromised for both types of scrapie in some countries.

**scrapie / surveillance / meta-analysis / evaluation**

## 1. INTRODUCTION

Regulation 999/2001 [1] required the implementation in 2002 of two surveys to improve the detection of scrapie throughout Europe: the abattoir (AS) and the fallen stock (FS) surveys. In general terms the AS samples sheep older than 18 months of age fit for human consumption at the abattoir [12]. Theoretically this is a random sample of the slaughtered sheep population. The FS targets dead-on-farm sheep reported by farmers. Animals must also be older than 18 months of age [6]. There is evidence that the risk of scrapie in the latter is larger than that in the AS[1].

---

[1] Wilesmith J.W., Ryan J., Del Rio Vilas V.J., Gubbins S., Summary of the results of scrapie surveillance in sheep in Great Britain, April-December 2003, Veterinary Laboratories Agency (2004) http://www.defra.gov.uk/animalh/bse/publications/reports/SheepSurvey2.pdf [consulted Feb. 2008].

---

\* Corresponding author: v.delriovilas@vla.defra.gsi.gov.uk. As a UK Civil Servant, his work is Crown copyright.

Regulation 999/2001 [1] establishes a set of methodological criteria that the European Union countries must follow as an attempt to standardise the surveys. These criteria should reduce the potential between-country heterogeneity, at least from a methodological point of view. This would facilitate direct comparison of the results between countries.

Despite the efforts to standardize the surveys across the EU, variability between countries, within the same year, may still occur due to differences in the risk of scrapie in the populations sampled by either survey. Differences between years, within countries, are likely to originate as a result of changes in the methodology of the surveys. These changes can occur as a result of modifications derived from new EU mandates, hence, affecting all countries, e.g. the increase in 2004 of the sample size of the fallen stock survey, or due to individual countries modifying their approach for any reason, e.g. the application of different screening tests.

Previous studies [7] assessed the presence of heterogeneity between the observed prevalence estimates of 18 EU countries by means of a meta-analysis of the log-odds ratios (FS/AS), adjusting for country-specific covariate information and the risk in the baseline population (AS), and showed a large residual variability indicating an inconsistent approach to the surveys across the EU. As a result of the remainder heterogeneity, the pooled odds-ratio (FS/AS) informing the efficacy of the FS relative to that of the AS could not be reported. The authors pointed out that a potential source of heterogeneity could have come from the joint modelling of data on both types of scrapie: classical and atypical. The separate analysis of the two types of scrapie merits further study. The inclusion of extra covariate information might also explain further the variability between countries. The EU annual report[2], from which previous studies [7] collected their data, did not provide

additional covariate information or enough detail to model the two types of scrapie separately.

An alternative approach to overcome the lack of covariate information, still under the general context of the meta-analysis, is to model the unobserved heterogeneity in our data. Countries could be grouped into clusters representing the underlying subpopulations relative to the risk of scrapie between the two surveys in each country. This cluster structure can be modelled by means of non-parametric mixture distributions [15]. Non-parametric approaches have been suggested as a way to avoid the potential limitations associated with assuming specific forms of parametric distributions, especially when dealing with small samples and subject to uncertainty about the number of subgroups/clusters in the population [5]. Modelling the unobserved heterogeneity in the data would facilitate comparisons between countries within clusters where pooled estimates, of the efficacy of the FS relative to that of the AS, can be reported. Ultimately, the identification of the discrepancies between countries may increase our understanding of the underlying country-specific characteristics monitored by the surveys.

This study complements and extends a previous work [7] in an attempt to explain further the processes underlying the scrapie surveillance figures across the EU. It does so by applying recent methodological applications [15]. Furthermore, we are now in position to inform our models with specific scrapie-type data. This will allow the modelling of the two types of scrapie separately and, indirectly, the assessment of the effect of the joint modelling on the overall heterogeneity in our previous models. The objective of this study is three fold: (i) to model the cluster structure of EU's active surveillance data on classical and atypical scrapie, (ii) to compare this structure between years to assess

---

[2] European Commission, Report on the monitoring and testing of ruminants for the presence of transmissible spongiform encephalopathy (TSE) in the EU in 2003, including the results

of the survey of the prion protein genotypes in sheep breeds, Brussels (2004), http://ec.europa.eu/food/food/biosafety/bse/annual_report_tse2003_en.pdf [consulted February 2008].

the standardisation of the surveillance efforts throughout the EU and (iii) to evaluate the performance of the two surveys in the detection of the two types of scrapie.

## 2. MATERIALS AND METHODS

### 2.1. Materials

Data on the number of sheep tested and confirmed by each surveillance source (AS and FS) were collected from the EU's annual reports on the monitoring of transmissible spongiform encephalopathy (TSE) in ruminants in 2003[2], 2004[3] and 2005[4]. Data from 18 European countries were selected in all three years (Tab. I).

Furthermore, data on the number of cases of atypical scrapie, for 2003, 2004 and 2005, by surveillance source and country were obtained directly from the EU Commission. Atypical figures by country and surveillance source for 2004 and 2005 are also shown in Table I. We used 2003 data on atypical scrapie only to assess the impact of the joint modelling of both diseases on the overall model's heterogeneity. For the rest of the analyses on atypical scrapie we used 2004 and 2005 data only. We dismissed 2003 data to allow for a consistent case-definition for atypical scrapie across the countries. In the case of Great Britain, atypical cases were firstly reported as "unconfirmed Biorad positives" (UBP)[1]. A similar misclassification was likely to occur in other countries in the early years of the active surveillance.

### 2.2. Methods

The comparison of some form of frequency ratio between the two surveys across the EU, under

---

[3] European Commission, Report on the monitoring and testing of ruminants for the presence of transmissible spongiform encephalopathy (TSE) in the EU in 2004, Brussels (2005), http://ec.europa.eu/food/food/biosafety/bse/annual_report_tse2004_en.pdf [consulted February 2008].
[4] European Commission, Report on the monitoring and testing of ruminants for the presence of transmissible spongiform encephalopathy (TSE) in the EU in 2005, Brussels (2006), http://ec.europa.eu/food/food/biosafety/bse/annual_report_tse2005_en.pdf [consulted February 2008].

the standard conditions that apply to the surveys' operations, appears as an adequate methodology to assess the comparability of the scrapie surveillance in Europe. This was the approach pursued by Bird [2]. More recently, meta-analysis techniques were applied to this goal [7]. Meta-analysis methodologies are well suited to: (1) demonstrate the occurrence of heterogeneity in the way scrapie surveillance is conducted across the EU, and (2) evaluate the performance of the two surveys in the detection of the two types of scrapie. The explanation of the heterogeneity between studies is a logical step and one of increasing importance when conducting meta-analyses [17, 19]. For the application of these techniques, we need some measure of effect: odds-ratios, risk-ratios or rate-ratios. This measure of effect not only allows the comparison between the efficacies of the two surveys in the detection of scrapie, but is also an indicator of the underlying prevalence in each country and the artefacts (presented in the form of covariates) that may affect our perception of it.

We start our analyses by defining the measure of effect to be compared between the EU countries under the meta-analysis approach. Data in Table I can be displayed in a $2 \times 2$ table and risk ratios $\theta_i = x_i^{FS} n_i^{AS} / x_i^{AS} n_i^{FS}$, where $x_i^{FS}(x_i^{AS})$ represent the number of cases of scrapie detected and $n_i^{FS}(n_i^{AS})$ the number of samples tested in the FS(AS), between the FS and the AS computed for each country $i$. This risk ratio indicates how likely is for a scrapie case to be detected by the FS relative to the AS.

To study the presence of overall heterogeneity between the countries we conducted a random-effects meta-analysis [9] of the $\theta_i$ for all $k$ countries and compute the Q statistic [10] that measures the overall variability in the meta-analysis. We did so for all three years of classical scrapie data. For atypical scrapie we performed the same analyses on 2005 data alone, to allow for the incorporation of year-specific covariates, and the pooled 2004–2005 dataset to increase the power of our analysis. The overall heterogeneity $\tau^2$ was estimated from the Mantel-Haenszel approach that has been suggested as more adequate for studies with zero events [18]. We also compute the $I^2$ statistic [14] for the quantification of the heterogeneity due to between-country variability. Galbraith plots [13] to visualize the country-specific contribution to the overall heterogeneity were produced as an aid to interpret our results.

V. J. Del Rio Vilas et al.

**Table I.** Data by type of scrapie, surveillance source and year. $x_{FS}$ shows the number of cases detected by the fallen stock survey, $n_{FS}$ the number of sheep sampled by the fallen stock survey, $x_{AS}$ shows the number of cases detected by the abattoir survey and $n_{AS}$ the number of sheep sampled by the abattoir survey.

| Countries | Classical scrapie | | | | | | | | | | | | Atypical scrapie | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| | 2003 | | | | 2004 | | | | 2005 | | | | 2004 | | 2005 | |
| | $x_{FS}$ | $n_{FS}$ | $x_{AS}$ | $n_{AS}$ | $x_{FS}$ | $n_{FS}$ | $x_{AS}$ | $n_{AS}$ | $x_{FS}$ | $n_{FS}$ | $x_{AS}$ | $n_{AS}$ | $x_{FS}$ | $x_{AS}$ | $x_{FS}$ | $x_{AS}$ |
| Belgium | 2 | 496 | 0 | 2 376 | 2 | 1 516 | 1 | 39 | 1 | 1 451 | 0 | 10 | 1 | 0 | 1 | 0 |
| Denmark | 0 | 1 320 | 0 | 871 | 0 | 5 253 | 0 | 91 | 0 | 4 295 | 0 | 97 | 0 | 0 | 0 | 0 |
| Germany | 13 | 48 629 | 9 | 20 116 | 42 | 65 488 | 1 | 15 628 | 18 | 29 550 | 8 | 14 894 | 0 | 0 | 0 | 0 |
| Greece | 13 | 793 | 49 | 22 613 | 17 | 2 098 | 4 | 6 508 | 34 | 1 597 | 13 | 4 484 | 0 | 0 | 0 | 0 |
| Spain | 7 | 12 950 | 13 | 49 940 | 7 | 10 799 | 7 | 15 051 | 17 | 14 881 | 14 | 14 274 | 1 | 0 | 1 | 1 |
| France | 27 | 18 989 | 23 | 44 687 | 22 | 12 125 | 11 | 12 456 | 37 | 22 411 | 3 | 12 246 | 2 | 9 | 5 | 8 |
| Ireland | 18 | 2 848 | 9 | 51 588 | 36 | 9 632 | 5 | 10 686 | 21 | 10 374 | 1 | 10 689 | 1 | 0 | 0 | 0 |
| Italy | 13 | 5 024 | 14 | 35 274 | 4 | 4 931 | 8 | 16 839 | 16 | 8 398 | 6 | 14 173 | 0 | 0 | 1 | 4 |
| Luxembourg | 0 | 244 | 0 | 213 | 0 | 299 | 0 | 125 | 0 | 428 | 0 | 238 | 0 | 0 | 0 | 0 |
| Netherlands | 6 | 4 000 | 45 | 21 140 | 26 | 10 137 | 13 | 8 949 | 22 | 10 085 | 13 | 8 910 | 0 | 0 | 1 | 1 |
| Austria | 0 | 3 255 | 0 | 4 225 | 0 | 2 052 | 0 | 446 | 0 | 4 180 | 0 | 116 | 0 | 0 | 0 | 0 |
| Portugal | 0 | 243 | 5 | 10 703 | 0 | 1 470 | 0 | 42 753 | 0 | 21 230 | 0 | 51 285 | 0 | 24 | 25 | 32 |
| Finland | 0 | 683 | 0 | 1 990 | 0 | 802 | 0 | 501 | 0 | 899 | 0 | 394 | 1 | 0 | 1 | 0 |
| Sweden | 0 | 2 849 | 0 | 5 177 | 0 | 2 985 | 0 | 166 | 0 | 3 239 | 0 | 1 | 2 | 0 | 1 | 0 |
| United Kingdom | 13 | 5 126 | 45 | 72 518 | 12 | 5 091 | 9 | 11 304 | 24 | 9 761 | 16 | 11 816 | 4 | 12 | 9 | 14 |
| Czech Republic | 0 | 2 528 | 1 | 426 | 2 | 896 | 0 | 55 | 1 | 360 | 0 | 35 | 0 | 0 | 0 | 0 |
| Slovakia | 1 | 214 | 1 | 3 924 | 10 | 661 | 5 | 1 155 | 7 | 2 365 | 2 | 250 | 0 | 0 | 0 | 0 |
| Norway | 0 | 3 367 | 0 | 33 524 | 0 | 3 367 | 0 | 10 462 | 0 | 3 615 | 0 | 10 889 | 4 | 8 | 2 | 1 |

The two types of heterogeneity that can be found in a meta-analysis have been described previously [15]: the effect heterogeneity resulting from the different interventions across the studies (countries in our case) and the baseline heterogeneity. The latter is a well-known and commonly reported source of variability in the meta-analysis literature [19, 21]. Baseline heterogeneity, in our setting, arises when the risks of scrapie observed in the abattoir survey vary in a significant manner across the EU countries. The existence of any form of heterogeneity in a significant amount would invalidate any pooled inference from the meta-analysis.

To estimate the $\theta_i$ under conditions of unobserved heterogeneity we applied the profile likelihood (PL) model. A comprehensive description of the methodology in the context of the meta-analysis is provided elsewhere [15]. Under conditions of unobserved heterogeneity instead of having a specific value $\theta_i$ for each country, we consider a distribution on the range of the possible values of $\theta$. This last distribution is estimated non-parametrically by a discrete mixture giving weight $q_j$ to the relative risks $\theta_j$. In this setting, the PL method is used to inform the parametric density $f(x_i|\theta_j)$, also called the mixture kernel, of the non-parametric mixture distribution

$$f(x_i|P) = \sum_{j=1}^{m} f(x_i|\theta_j)q_j$$

where $x_i$ contains for country $i$ the following data in vector form $(x_i^{AS}, n_i^{AS}, x_i^{FS}, n_i^{FS})$, $P$ is a discrete mixing distribution $P = \begin{pmatrix} \theta_1 \ \theta_2 \ ... \ \theta_m \\ q_1 \ q_2 \ ... \ q_m \end{pmatrix}$ with positive weights $q_j > 0$ given to component relative risk $\theta_j$, and $\sum_{j=1}^{m} q_j = 1$. The number of components or clusters of countries are $m$. Normally $m$ will be small in comparison with the number of studies, countries in our case. If $m = 1$ this is the homogeneity situation. In a situation of extreme heterogeneity, the number of clusters will equal that of the studies/countries.

Now, we find that the density function for the PL approach used as kernel in the mixture distribution is

$$f_{PL}(x_i|\theta_i) = e^{-x_i} \frac{\left(\frac{n_i^{FS}\theta_i x_i}{n_i^{AS}+n_i^{FS}\theta_i}\right)^{x_i^{FS}} \left(\frac{n_i^{AS}x_i}{n_i^{AS}+n_i^{FS}\theta_i}\right)^{x_i^{AS}}}{x_i^{FS}! \, x_i^{AS}!}$$

It is important to note that the PL approach integrates the baseline risk ($p_i^{AS} = x_i^{FS} n_i^{AS}$), a nuisance parameter, into the occurring likelihood.

This makes this approach the preferred one in situations of baseline heterogeneity [15].

We finally need to estimate the parameters of the mixing distribution $P$. We did so by means of maximum likelihood estimation via the expectation-maximisation (EM) algorithm and the introduction of a latent indicator variable $z_{ij}$ that takes value 1 if country $i$ belong to component $j$ and 0 otherwise. The E-step of the EM algorithm has the form

$$e_{ij} = E(z_{ij}|x_i, q_j, \theta_j) = \frac{q_j f(x_i|\theta_j)}{\sum_{l=1}^{m} q_l f(x_i|\theta_l)}.$$

The classification of the countries into clusters is possible after the estimation of the parameters $(\theta_j, q_j)$ because $e_{ij}$ can be read as the posterior probability of country $i$ belonging to cluster $j$.

The number of clusters of countries ($m$) is unknown and must be estimated. In this study we consider the Bayesian information criterion (BIC) for the choice of the mixture model. The BIC is defined as $BIC = 2l(P) - d\log(k)$ with $d$, the number of estimated parameters, $d = (2m - 1)$ for the PL model, $k$ the number of countries and $l$ the maximized value of the profile likelihood function for the estimated model. We will choose the mixture model with the largest BIC value.

The extension of our analyses to allow comparisons in time can be done either by conducting the three sets of analysis as above, one for each year, to account only for the between-country heterogeneity or alternatively, modelling the three years together, hence accounting simultaneously for the time and between-country heterogeneity. The latter approach presents a series of methodological advantages [3, 4]. We pooled the three years of data on classical scrapie to assess time-clusters effects. The time-country analysis produces a unique cluster structure that can evaluate better the performance of a country over the years.

### 2.3. Covariate information

Our models can be extended to incorporate country-specific covariates informing systematic variability. The incorporation of covariates will allow the investigation of the causes of heterogeneity between countries. Previous studies [7] found a significant positive association between the proportion of the population sampled by the FS and the outcome under study, the logOR of the FS over the AS, in EU scrapie surveillance data for 2003. In the present study, we incorporate the same variable,

*repreFS* $= n_i^{FS}/N_i$ where $N_i$ is the country-specific adult sheep population for the year 2004[3]. The proportion of the adult sheep population sampled by the AS was not incorporated in the analyses. Previous works [7] did not find any relationship with the outcome.

Furthermore, we extended our models to incorporate the time effect as a covariate. Year was included as a categorical variable with 2003 as baseline in the three-year pooled dataset on classical scrapie.

In this setting we assumed that the covariate effects were constant across the mixture components, that is, that the slopes of the effect of the covariates were parallel in all components of the mixture. In essence we incorporate covariates in our models by means of a modification of the generalized linear model where the systematic component is expressed as a linear predictor $\eta$

$$\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip}$$

where $\beta_0, \beta_1, \ldots, \beta_p$ are the parameters of the model to be estimated by maximizing the PL and $x_{il}$ the value of the *l*th covariate for the *i*th country. A detailed description of the PL approach with covariates is provided by other authors[5]. Note that $\eta_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip}$ and $\theta_{ij}$, the relative risk of country *i* in cluster *j*, are connected via the log-link function $\exp(\eta_{ij}) = \theta_{ij}$. In other words, we have combined unobserved and observed heterogeneity in such a way that observed heterogeneity is modelled by means of a log-linear model and unobserved heterogeneity is captured via mixing in the intercept of the log-linear model.

The previous calculations, under the PL approach, were performed with the software CAMAP (Computer assisted analysis of meta-analysis type data with binomial outcome using the profile likelihood model of mixtures) developed by two of the authors (Böhning and Kuhnert).

## 3. RESULTS

Six countries in 2003 and seven in 2004 and 2005 were dropped from the analysis of

---

[5] Rattanasiri S., Modelling of covariate information in multi-centre studies with binary outcome using profile likelihood (Dissertation), Freie Universität Berlin, Berlin (2006), http://www.diss.fu-berlin.de/2006/344/indexe.html [consulted February 2008].

the classical scrapie data due to the presence of dual zeros in both surveillance sources. The countries analysed are shown in Table II. Continuity corrections (0.5) were added to allow the identifiability of the risk ratios (RR) for those countries with a zero value in either survey (Tab. II). The Q statistic was significant for all three years (Chi-squared: 87.83, d.f. 11, *p*-value < 0.0001 in 2003, Chi-squared: 33.03, d.f. 10, *p*-value < 0.0001 in 2004 and Chi-squared: 47.33, d.f. 10, *p*-value < 0.0001 in 2005) indicating the presence of heterogeneity in our datasets. Most of this variability was due to between-country heterogeneity as shown by the $I^2$ statistic: 87% (95% confidence intervals (CI): 80; 92) in 2003, 70% (95% CI: 44; 94) in 2004 and 79% (95% CI: 63; 88) in 2005. The Mantel-Haenszel estimator of $\tau^2$ shows a decline in the residual heterogeneity of the models from the large values in 2003 (1.41) to 0.66 and 0.85 in 2004 and 2005 respectively. The presence of residual heterogeneity invalidates the reporting of the pooled OR and justifies the modelling of the cluster structure in our dataset. Nevertheless, there is a clear decrease in the values of the pooled $\theta_{CLASS}$ across the years: 3.46 (95% CI: 1.57; 7.59) in 2003, 2.55 (95% CI: 1.39; 4.59) in 2004 and 1.97 (95% CI: 1.01; 3.85) in 2005. In 2005, the pooled $\theta_{CLASS}$ was just significant. The two surveys appear to get closer in their efficacy detecting classical scrapie.

Modelling the cluster structure of our data, without covariates, supported the results from the traditional meta-analyses for each individual year. The advantage now is that we can identify those countries that, belonging to the same cluster, show similar ratios and, consequently, can be compared with the confidence that they are reflecting similar underlying processes. The model with the largest BIC value in 2003 returned three clusters of countries. This compared with the models in 2004 and 2005 with two clusters each. Country-specific RR are shown in Table II for all three years. Recall here that the larger the number of clusters, the greater the heterogeneity. Furthermore, as the PL method integrates the baseline parameter

**Table II.** Cluster structure by year. Classical scrapie. RR shows the risk ratio for each country and year.

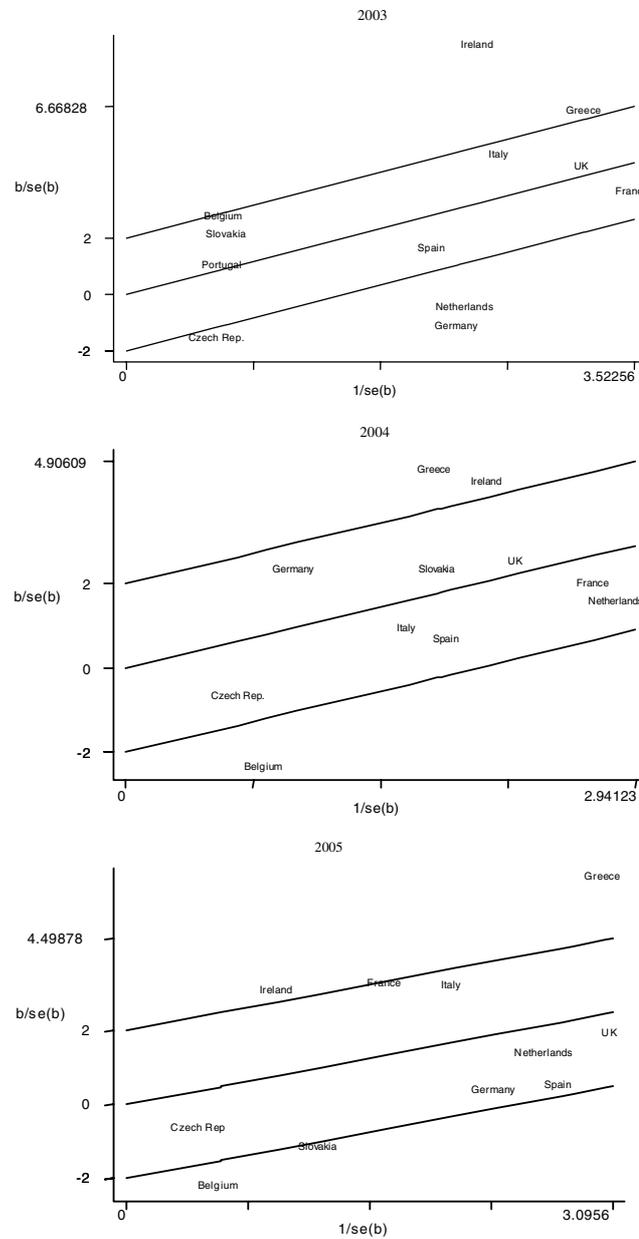| Countries | 2003 | | | 2004 | | | 2005 | | |
|---|---|---|---|---|---|---|---|---|---|
| | Cluster | RR | 95% CI | Cluster | RR | 95% CI | Cluster | RR | 95% CI |
| Belgium | 2 | 4.23 | 3.1; 5.78 | 1 | 2.06 | 1.41; 3 | 1 | 1.36 | 0.94; 1.95 |
| Germany | 3 | 0.65 | 0.34; 1.26 | 2 | 9.25 | 4.39; 19.52 | 1 | 1.36 | 0.94; 1.95 |
| Greece | 2 | 4.23 | 3.1; 5.78 | 2 | 9.25 | 4.39; 19.52 | 2 | 7.09 | 4.43; 11.36 |
| Spain | 2 | 4.23 | 3.1; 5.78 | 1 | 2.06 | 1.41; 3 | 1 | 1.36 | 0.94; 1.95 |
| France | 2 | 4.23 | 3.1; 5.78 | 1 | 2.06 | 1.41; 3 | 2 | 7.09 | 4.43; 11.36 |
| Ireland | 1 | 35.15 | 15.88; 77.78 | 2 | 9.25 | 4.39; 19.52 | 2 | 7.09 | 4.43; 11.36 |
| Italy | 2 | 4.23 | 3.1; 5.78 | 1 | 2.06 | 1.41; 3 | 2 | 7.09 | 4.43; 11.36 |
| Netherlands | 3 | 0.65 | 0.34; 1.26 | 1 | 2.06 | 1.41; 3 | 1 | 1.36 | 0.94; 1.95 |
| United Kingdom | 2 | 4.23 | 3.1; 5.78 | 1 | 2.06 | 1.41; 3 | 1 | 1.36 | 0.94; 1.95 |
| Czech Republic | 2 | 4.23 | 3.1; 5.78 | 1 | 2.06 | 1.41; 3 | 1 | 1.36 | 0.94; 1.95 |
| Slovakia | 3 | 0.65 | 0.34; 1.26 | 1 | 2.06 | 1.41; 3 | 1 | 1.36 | 0.94; 1.95 |
| Portugal | 2 | 4.23 | 3.1; 5.78 | – | – | – | – | – | – |

into the estimation of the parameter of interest (the treatment effect or, in our case, the FS), the heterogeneity observed in our results derives only from that of the heterogeneous application of the FS across countries.

Galbraith plots are useful tools to visualize the countries' contribution to the overall heterogeneity. Figure 1 shows the plots for all three years. As expected, the resemblance between the position of the countries in the plot and the $\theta_i$ shown in Table II from our crude PL models are evident. In 2003, Ireland appears alone at the top of the plot. This matches the cluster structure returned by the PL model in which Ireland appears alone. Germany, the Netherlands and the Czech Republic appeared grouped in the same cluster from the PL model. These three countries appear well separated from the rest in the Galbraith plot. In 2004, Germany, Greece and Ireland are the countries that contributed the most to the heterogeneity of the model. These countries appeared grouped in the same cluster from the PL approach. In 2005, Ireland, Greece, France and Italy appear at the top of the Galbraith plot. This matches our results shown in Table II. Across the years, Ireland and Greece appear to contribute consistently to the overall heterogeneity.

The results from the inclusion of the covariate repreFS in our models are shown in Table III. In 2003, the model with the largest BIC value corresponded to that with three clusters of countries. It is important to note at this point another advantage of modelling the cluster structure of the data compared to the traditional meta-regression analysis. We can now observe the cluster-specific effects of the covariate. This is obvious for 2003 and 2005 (Tab. III) where one of the components show a non-significant effect of repreFS on the cluster-specific RR. For 2004, repreFS did not appear significantly associated with the RR in any of the two clusters. It is important to note that, for all the clusters where repreFS had a significant effect, the sign of the coefficient was negative.

After the inclusion of repreFS, six of the eleven countries in 2005 showed RR not significantly different from 1, that is, their 95% confidence intervals included zero: Belgium (RR = 0.32; 95% CI: 0.002; 61.41), Germany (RR = 0.79; 95% CI: 0.4; 1.64), Spain (RR = 1.42; 95% CI: 0.88; 2.29), UK (RR = 1.47; 95% CI: 0.91; 2.4), Czech Republic (RR = 0.95; 95% CI: 0.21; 4.28), Slovakia (RR = 0.48; 95% CI: 0.01; 22.18). In other words, there was no significant improvement in the detection of classical scrapie by the FS relative to the AS. This compares with three countries in 2003 (Germany (RR = 0.56; 95% CI: 0.24; 1.32), Netherlands (RR = 0.84; 95% CI: 0.37; 1.93), Czech Republic (RR = 0.08; 95% CI: 0.002; 3.06)) and just one in 2004 (Germany (RR = 3.75; 95% CI: 0.53; 26.67) with

**Figure 1.** Galbraith plots by year for classical scrapie only (2003, 2004 and 2005). The log-risk ratios (b) divided by the standard errors for those EU countries with no dual zeros plotted against the reciprocal of the standard errors (horizontal axis). Solid lines represent the unweighted regression line constrained at 0 with a slope equal to the overall log-risk ratios of a fixed effects meta-analysis on our data and its 95% limits on both sides. The position of the countries in the $y$-axis indicates their contribution to the Q statistic for heterogeneity. The position of the countries on the $x$-axis indicates the weight of each country in the meta-analysis.

**Table III.** Cluster structure by year after the incorporation of the covariate "repreFS" (proportion of the adult sheep population sampled by the fallen stock survey in each country). Classical scrapie. Component shows the number of clusters of counties in each year. The column "Countries" shows the number of countries grouped under the same component of the mixture. The last four columns show the terms, intercept and covariate, regression-like coefficients, standard errors and level of significance of the generalized linear model used to estimate the relative risk in the presence of covariate information.

| Year | Component | Countries | Term | Coefficient | s.e. | $p$-value |
|------|-----------|-----------|------|-------------|------|-----------|
| 2003 | 1 | 3 | Intercept | 3.7 | 0.42 | < 0.0001 |
|      |   |   | repreFS | –0.19 | 0.03 | < 0.0001 |
|      | 2 | 8 | Intercept | 1.63 | 0.2 | < 0.0001 |
|      |   |   | repreFS | –0.19 | 0.09 | 0.016 |
|      | 3 | 1 | Intercept | 0.35 | 1.42 | 0.4 |
|      |   |   | repreFS | –0.19 | 0.49 | 0.35 |
| 2004 | 1 | 9 | Intercept | 0.67 | 0.22 | 0.001 |
|      |   |   | repreFS | 0.02 | 0.03 | 0.28 |
|      | 2 | 2 | Intercept | 2.19 | 0.42 | < 0.0001 |
|      |   |   | repreFS | 0.02 | 0.09 | 0.41 |
| 2005 | 1 | 6 | Intercept | 2.01 | 0.24 | < 0.0001 |
|      |   |   | repreFS | –0.16 | 0.03 | < 0.0001 |
|      | 2 | 5 | Intercept | 0.45 | 0.29 | 0.06 |
|      |   |   | repreFS | –0.16 | 0.29 | 0.29 |

similar non-significant results. To appreciate the effects of repreFS on the RR, the results above may be compared to those in Table II.

Table IV shows the results of modelling the countries and years together (33 observations after the removal of Portugal from 2003) adjusted for repreFS. This covariate appeared significantly associated with the RR ($\beta = -0.04$, $p$-value: 0.01) for the two clusters returned by the most parsimonious model. Across the years, the countries show differing patterns. Germany and the Czech Republic showed a consistent poor performance of the FS relative to the AS in the detection of classical scrapie. The UK, Netherlands and Spain appeared clustered together throughout the three years. The FS for these three countries showed a consistent moderate performance relative to the AS. Ireland and Greece appeared grouped in the same clusters consistently across the years but in this case showing a better performance of the FS. Belgium, Italy and Slovakia showed the same clustering pattern with a clear decrease in the performance of the FS relative to the AS across the years.

This is in contrast with France where the FS appears to increase its efficacy relative to the AS throughout the three-year period.

Overall, the joint modelling of country and time heterogeneity does not seem to indicate a decreasing trend in the RR of classical scrapie along the years. This is also supported by the results after the incorporation of year as covariate. Neither 2004 nor 2005 showed a significant effect when compared with the 2003 baseline. This interpretation would differ from that obtained from the comparison of the year-specific outputs from both the PL approach and the traditional meta-analysis.

The analysis of the 2005 atypical data, without repreFS, showed that all ten countries are clustered together. This was the model with the largest BIC. The pooled RR for this dataset was 1.12 (95% CI: 0.75; 1.68) indicating that there was non-significant difference between the two surveys with regard to the detection of atypical scrapie. When we modelled 2004 and 2005 data together, the largest BIC corresponded to the model with two components (Tab. V). In this last case, the

**Table IV.** Cluster structure of the pooled data (33 observations). RR shows the risk ratio for each country and year adjusted for the covariate "repreFS" (proportion of adult sheep sampled by the fallen stock survey).

| Countries | Year | Component | RR | 95% CI |
|---|---|---|---|---|
| Belgium | 2003 | 1 | 8.88 | 6.47; 12.16 |
| Belgium | 2004 | 2 | 1.32 | 0.98; 1.79 |
| Belgium | 2005 | 2 | 1.35 | 1.01; 1.8 |
| Germany | 2003 | 2 | 0.78 | 0.39; 1.56 |
| Germany | 2004 | 1 | 2.78 | 0.91; 8.55 |
| Germany | 2005 | 2 | 1.14 | 0.77; 1.7 |
| Greece | 2003 | 1 | 10.21 | 7.32; 14.24 |
| Greece | 2004 | 1 | 10.15 | 7.19; 14.13 |
| Greece | 2005 | 1 | 10.18 | 7.31; 14.18 |
| Spain | 2003 | 2 | 2.01 | 1.57; 2.59 |
| Spain | 2004 | 2 | 2.02 | 1.57; 2.58 |
| Spain | 2005 | 2 | 2.01 | 1.56; 2.58 |
| France | 2003 | 2 | 1.88 | 1.5; 2.36 |
| France | 2004 | 2 | 1.95 | 1.54; 2.46 |
| France | 2005 | 1 | 9.21 | 6.72; 12.62 |
| Ireland | 2003 | 1 | 9.98 | 7.2; 13.83 |
| Ireland | 2004 | 1 | 9.37 | 6.83; 12.85 |
| Ireland | 2005 | 1 | 9.3 | 6.79; 12.75 |
| Italy | 2003 | 1 | 9.98 | 7.21; 13.84 |
| Italy | 2004 | 2 | 2.01 | 1.58; 2.58 |
| Italy | 2005 | 2 | 1.97 | 1.55; 2.51 |
| Netherlands | 2003 | 2 | 1.83 | 1.48; 2.28 |
| Netherlands | 2004 | 2 | 1.53 | 1.22; 1.92 |
| Netherlands | 2005 | 2 | 1.53 | 1.22; 1.93 |
| UK | 2003 | 2 | 2.04 | 1.58; 2.65 |
| UK | 2004 | 2 | 2.04 | 1.58; 2.65 |
| UK | 2005 | 2 | 2.03 | 1.57; 2.61 |
| Czech Republic | 2003 | 2 | 0.81 | 0.41; 1.57 |
| Czech Republic | 2004 | 2 | 1.48 | 1.16; 1.88 |
| Czech Republic | 2005 | 2 | 1.8 | 1.46; 2.24 |
| Slovakia | 2003 | 1 | 9.96 | 7.19; 13.8 |
| Slovakia | 2004 | 2 | 1.89 | 1.51; 2.37 |
| Slovakia | 2005 | 2 | 1.51 | 1.19; 1.91 |

FS of Spain, Ireland, Norway, Portugal and Finland, in cluster 2 (pooled RR = 1.77, 95% CI: 1.1; 2.84), detected more atypical scrapie than the AS.

The inclusion of repreFS when we modelled the 2005 atypical data showed the significant effect of this covariate. In this case the effect was positive ($\beta = 0.14$, $p$-value = 0.03). The model with the largest BIC grouped all ten countries in a single cluster. The significant effect of repreFS leads to country-specific

$\theta_{iATYP}$ (Tab. V) although their 95% confidence intervals, for all countries, overlap.

## 4. DISCUSSION

The approach applied in the present study, modelling the cluster structure across the EU countries, acknowledges the existence of heterogeneity in the surveillance figures throughout Europe. This was obvious from the values of the Q statistic and those of

**Table V.** Cluster structure by year. Atypical scrapie. RR shows the risk ratio for each country. RR, for 2005 only, are adjusted for the proportion of the adult sheep population sampled by the fallen stock survey (repreFS).

| Countries | 2005 (with repreFS) | | | 2004–2005 (no repreFS) | | |
|---|---|---|---|---|---|---|
| | Cluster | RR | 95% CI | Cluster | RR | 95% CI |
| Belgium | 1 | 2.6 | 0.9; 7.1 | 2 | 0.55 | 0.31; 0.97 |
| Spain | 1 | 0.71 | 0.36; 1.36 | 1 | 1.77 | 1.1; 2.84 |
| France | 1 | 0.92 | 0.58; 1.45 | 2 | 0.55 | 0.31; 0.97 |
| Italy | 1 | 0.75 | 0.41; 1.37 | 2 | 0.55 | 0.31; 0.97 |
| Netherlands | 1 | 1.71 | 0.93; 3.13 | 2 | 0.55 | 0.31; 0.97 |
| Portugal | 1 | 1.49 | 0.9; 2.47 | 1 | 1.77 | 1.1; 2.84 |
| Finland | 1 | 3.71 | 0.94; 14.7 | 1 | 1.77 | 1.1; 2.84 |
| Sweden | 1 | 1.74 | 0.94; 3.24 | 2 | 0.55 | 0.31; 0.97 |
| United Kingdom | 1 | 0.68 | 0.34; 1.36 | 2 | 0.55 | 0.31; 0.97 |
| Norway | 1 | 0.98 | 0.98; 1.5 | 1 | 1.77 | 1.1; 2.84 |

the residual heterogeneity for all three years. This heterogeneity would stop the reporting of any pooled estimate, for all EU countries under study, on the ability of the fallen stock survey to detect scrapie vs that of the abattoir survey. At least, this was the case for classical scrapie. For atypical scrapie, our results show that all ten countries with sufficient data in 2005 appear under the same cluster. Although with different estimates of the RR due to the effects of the covariate, there appears not to be a clear benefit in using the FS to detect atypical scrapie relative to the AS. For classical scrapie, the declining values of the pooled RR from the traditional meta-analyses seem to lead to a similar interpretation. Although with substantial residual heterogeneity, the decreasing values of the RR show how the two surveys are getting closer in their ability to detect classical scrapie. In other words, and keeping other things equal (e.g. tests used in both surveys), the populations sampled by the two surveys are becoming similar in their risks of scrapie. Such an event would stop the gains in efficacy expected from the FS in terms of detecting classical scrapie. The opposite explanation, an increase in the risk of scrapie among the slaughtered population seems less likely but still possible (see below). The AS remains, in general, a random survey or, at least, one where interventions (e.g. increased sample sizes) are less likely to impact on the observed estimates. This is so because of the

subsidiary nature of the AS with regard to the slaughter trade of sheep. Modifications in the AS would rarely affect the slaughter trade. On the other hand, modifications in the FS are much more likely to reflect on the prevalence estimates as farmers can change their patterns due to the voluntary nature of the survey. Interventions at different levels can alter the reliability of the samples as reported by previous studies [7], e.g. in the case of Germany where younger animals were sampled by the FS.

Known to the authors are the costs associated to both surveys in the UK (data not shown). For this particular country, the cost of the FS, per sample, is several times that of the AS. This is explained by the expensive logistic network required for the prompt collection of the samples anywhere in the country. This is much less in the case of the AS where the animals are delivered to the point of sampling at no expense for the survey. The increased costs per sample together with a moderate efficacy in detecting the event require joint consideration in the evaluation of the FS.

From the traditional meta-analysis approach we would have concluded that the heterogeneity across the years dropped and the RR for classical scrapie decreased. The pooled analysis of all years returned a dissimilar interpretation, with no significant differences between the RR across time and, in general, most countries showing RR greater than one

in all three years. In other words, the FS remained relatively better than the AS in detecting classical scrapie although less than the expected ten times better, by analogy with BSE, as anticipated by the European Scientific Steering Committee (ESSC)[6]. For some countries, Greece and Ireland, the RR were close to those expected values (Tab. IV). However, as we observed from the Galbraith plots, these two countries contributed the most to the overall heterogeneity. The norm, represented by the remaining nine countries, did not return such large RR. The clustering of the countries in two components is still a strong indicator of the presence of heterogeneity across the EU.

Whilst it seems that the FS is less efficient than initially expected, this may have nothing to do with the methodology of the survey per se but with changes in the underlying characteristics of the targeted populations or changes in the biology of the disease. We might be detecting a plausible increase of sub-clinical infections, detected with greater ability by the AS, or a decrease in clinical apparent disease with the result of death, as detected by the FS (assuming that the majority of the cases detected by this source die of scrapie). Trying to get to the root of the problem, a change in the genotype distribution, the breed distribution or the age distribution of the population might well explain our results across the years. Breeding programmes, genotyping schemes and changes in the subsidies to farmers, perhaps explaining changes in the age of animals sent for slaughter, may be the primary causes. As much as these interventions may have some effect on the observed prevalence estimates, and hence on our ratio measures, this is likely to be felt in the long term in the

case of the breeding programmes. On the other hand, the effects of any change in the age of the slaughtered population would be felt almost immediately. For example, if younger animals were sent for slaughter this might well result in an increase in the prevalence of classical scrapie detected by the AS.

This study did not classify the countries based on their scrapie situation: whether there is scrapie in the country and in what amount. We can carry out comparisons between countries even if the prevalence levels are disparate. In a homogenous situation, for countries A and B with very different prevalence levels, for example 1% and 10% respectively, the ratio FS/AS would be the same.

Previous studies [11] have stated the importance of standardised frameworks for the comparison of surveillance data. Trying to standardise the way the surveys are conducted across the EU should be a priority. This would allow informed comparisons between countries. Our year-specific results show that this seems to be the trend for classical scrapie with a reduction in the number of clusters of countries across the years. Ideally, this would be the result of a greater harmonization of surveillance procedures across the EU. A gradual homogenization of the populations across countries as a result of EU-wide breeding programmes would be an alternative explanation. The latter could mask variations in the implementation of the surveys and still show apparent gains in homogeneity between the countries. A questionable gain if it is at the expense of the natural heterogeneity in the populations. The scarce data available from the EU Commission reports does not allow the identification of the specific sources, methodological vs. population-based, behind this reduction in the heterogeneity. The same techniques shown here could be applied to explain changes in the observed ratios or prevalence estimates across time adjusted for population, e.g. genotype data, and methodological characteristics, e.g. tests.

As important as the study of the heterogeneity between countries is the assessment of the consistency within countries in their surveillance efforts across the years. Countries

---

[6] Scientific Steering Committee: Opinion on requirements for statistically authoritative BSE/TSE surveys, Adopted by the Scientific Steering Committee at its meeting of 29–30 November 2001, Brussels, European Commission, Health & Consumer Protection Directorate-General (2001), http://ec.europa.eu/food/fs/sc/ssc/out238_en.pdf [consulted February 2008].

clustered together would appear to share similar characteristics of the targeted populations or conduct the surveys in a similar way. Our results showed large variations in the values of $\theta_{iCLASS}$ throughout time for some countries. Belgium, probably due to the reduced number of samples, France, Italy and Slovakia appear with varying estimates. We have already mentioned that the most likely explanation originates from changes in the operation of the surveys. The lack of consistency in the implementation of the surveys will, no doubt, affect the study of trends and of the results from the regular evaluation of the surveillance network.

We have focused our analyses on the active component of the surveillance of scrapie, the two surveys, where we can actively modify some of the operational parameters with relative ease in order to improve the attributes of the surveillance system [16]. Passive surveillance figures, namely the reporting of clinical suspect sheep by farmers, were also available from the EU reports and, hence, could have been included in our analyses. We decided against it in the pursuit of efficiency: passive surveillance cannot be modified but influenced making the return from our improvement efforts less certain. This would follow the rationale on the appropriateness of screening programmes by which screening is only worthwhile for those disorders for which effective interventions exist. Furthermore, the analysis of the passive surveillance would probably show large heterogeneity between the countries rising from the unrestricted scope of the source with no uniform set of operational criteria as in the case of the surveys.

As a by-product of our analysis we could assess the impact of the joint modelling of the two types of scrapie. This was the approach pursued in previous studies [7]. In general, we did not find large differences when we modelled the two types of scrapie together in comparison with the results from the modelling of classical scrapie alone. It appears that the number of atypical cases did not bear enough weight to change our results, as the number of clusters remained the same for the three years.

To inform completely our models we would need detailed data on the tests applied in both surveys by all countries. Recent studies [20] have shown the effects of the screening test of choice on the surveillance results observed in Great Britain. This variable was incorporated, after an extensive and time-consuming request to all EU National Institutes, in previous analyses on 2003 data [7]. Despite their efforts, the authors argued that they did not have enough data to model all the potential heterogeneity from this variable. Data on this variable was not readily available from the EU reports for the present analyses. Nor it was possible to obtain it from the Commission due to confidentially issues. Future work would benefit from a centralized and standardised collection of data on this variable.

The homogenous picture in 2005 for atypical scrapie, with only one cluster, contrasts with that of classical with two clusters. The question of why the surveillance appears less heterogeneous for atypical scrapie than for classical is immediate. Without more data on atypical scrapie, for more years and larger numbers, we cannot assess the direction of the effects for this disease.

Two clusters with opposite signs in terms of the FS's ability to detect atypical scrapie were well defined from the pooled analysis of 2004 and 2005 data. Perhaps our results from this pooled analysis provide a fairer, based on the larger numbers, image of the FS's performance with regard to atypical scrapie. If so, the FS would also show a heterogeneous performance for this type of scrapie too.

The value of the coefficient $\beta$ (0.14) for repreFS when modelling the 2005 atypical dataset indicates a positive correlation between the covariate and our outcome. In other words, the larger the proportion of the adult sheep population sampled by the FS the greater the efficacy of this survey to detect atypical scrapie relative to the AS. This is different from classical scrapie. Pressure upon the farmers to submit all the dead-on-farm animals may alter the composition of the FS sample. The interventional nature of the survey makes it prone to this form of bias. For unforeseeable financial reasons, subsidized collection

schemes of fallen stock in combination with a decrease in subsidies per head might have increased the numbers of management culls reported to the FS. The effects of such rise might cause a drop in the ability of the FS to detect classical scrapie if, for example, older animals were included in the sample (the risk of classical scrapie appears small for older age-groups [8]). Inversely, a relative increase of older animals might increase the frequency of atypical scrapie detected by this survey.

The PL approach allowed us to expand further our efforts to inform the standardisation of the scrapie surveillance across the EU. In the absence of relevant covariates to inform our models, e.g. genotype and screening tests, the PL method allowed meaningful comparisons between countries within clusters. Furthermore, the integration of the baseline parameter in the profile likelihood, particularly in the presence of very likely strong baseline heterogeneity due to the observational nature of the data, is appealing and eases the interpretation of our findings [15]. Differences observed between countries are solely due to the heterogeneity in the FS.

As pointed out by a referee, the sources of heterogeneity between countries are numerous. In general, this heterogeneity can be grouped in population-based sources, related to, for example, country-specific farming practices, and methodological sources, related to the operation of the surveys. We did not try, at any point, to identify all the sources of heterogeneity as they are too many. Furthermore, there is no data or any other form of evidence available to inform many of them for each country; for example, whether remote areas are covered or not. Our modest effort tries to inform as many sources of heterogeneity as possible in the form of covariates. The many other potential sources of heterogeneity that cannot be put in the form of covariates constitute the remainder heterogeneity that leads to the grouping of countries under different clusters. Whether, for example, clinical cases are not reported to the passive surveillance in some countries and end up in either survey, or different tests are used across the EU

with different sensitivities for either type of scrapie, they contribute to the unobserved heterogeneity that is specifically targeted in this study. Had we had data on these potentially important covariates our conclusions might have been different.

Other measures of effect could have been chosen: the odds ratio or the risk difference. Previous studies used the former [7] due to methodological advantages. For similar reasons, related to the current design of the software CAMAP, we used risk ratios in this study.

It would appear that the FS, due to its heterogeneity across countries, propensity to suffer from biases, high costs and less than expected efficacy to detect the event, might have lost its appeal as an efficient source for the detection of scrapie. We have shown that for many countries the FS failed to detect more classical scrapie than the AS, particularly in 2005. The poorer performance of the FS for atypical scrapie was also obvious from our results. Despite this, the FS would still work as a valid source if we could stop the reported selection biases and make it less expensive in its operation. The former can only be avoided with random schemes, much like the AS. The feasibility of randomly testing animals collected by national fallen stock schemes, like the National Fallen Stock Scheme Company in Great Britain, may merit study. The benefits of monitoring the different subpopulations are all too obvious to recommend, or allow via the reduction of its usefulness due to flawed designs, the elimination of an extra source of surveillance.

## REFERENCES

[1] Anonymous, Commission regulation (EC) No. 1248/2001 of 22 June 2001 amending annexes III, X and XI to regulation (EC) No 999/2001 of the European Parliament and of the Council as regards epidemio-surveillance and testing of transmissible spongiform encephalopathies, Off. J. Eur. Communities (2001) 44:12–22.

[2] Bird S.M., European Union's rapid TSE testing in adult cattle and sheep: implementation and results in 2001 and 2002, Stat. Methods Med. Res. (2003) 12:261–278.

[3] Böhning D., Dietz E., Schlattmann P., Space-time mixture modelling of public health data, Stat. Med. (2000) 19:2333–2344.

[4] Böhning D., Empirical Bayes estimators and non-parametric mixture models for space and time-space disease mapping and surveillance, Environmetrics (2003) 14:431–451.

[5] Congdon P., Bayesian statistical modelling, Wiley Series in Probability and Statistics, John Wiley & Sons, Chichester, 2001.

[6] Del Rio Vilas V.J., Ryan J., Elliott H.G., Tongue S.C., Wilesmith J.W., Prevalence of scrapie in sheep: results from fallen stock surveys in Great Britain in 2002 and 2003, Vet. Rec. (2005) 157:744–745.

[7] Del Rio Vilas V.J., Hopp P., Nunes T., Ru G., Sivam S.K., Ortiz-Pelaez A., Explaining the heterogeneous scrapie surveillance figures across Europe: a meta-regression approach, BMC Vet. Res. (2007) 28:3–13.

[8] Del Rio Vilas V.J., Guitian J., Pfeiffer D.U., Wilesmith J.W., Analysis of data from the passive surveillance of scrapie in Great Britain between 1993 and 2002, Vet. Rec. (2006) 159:799–804.

[9] DerSimonian R., Laird N., Meta-analysis in clinical trials, Control Clin. Trials (1986) 7:177–188.

[10] Dohoo I., Martin W., Stryhn H., Veterinary epidemiologic research, AVC Inc., Charlottetown, Canada, 2003.

[11] Donnelly C.A., Ferguson N.M., Ghani A.C., Anderson R.M., Implications of BSE infection screening data for the scale of the British BSE epidemic and current European infection levels, Proc. Biol. Sci. (2002) 269:2179–2190.

[12] Elliott H., Gubbins S., Ryan J., Ryder S., Tongue S., Watkins G., Wilesmith J.W., Prevalence of scrapie in sheep in Great Britain estimated from abattoir surveys during 2002 and 2003, Vet. Rec. (2005) 157:418–419.

[13] Galbraith R.F., A note on graphical presentation of estimated odds ratios from several clinical trials, Stat. Med. (1988) 7:889–894.

[14] Higgins J.P., Thompson S.G., Quantifying heterogeneity in a meta-analysis, Stat. Med. (2002) 21:1539–1558.

[15] Kuhnert R., Böhning D.A., Comparison of three different models for estimate relative risk in meta-analysis of clinical trials under unobserved heterogeneity, Stat. Med. (2007) 26:2277–2296.

[16] Romaguera R.A., German R.R., Klaucke D.N., Evaluating public health surveillance, in: Teutsch S.M., Churchill R.E. (Eds.), Principles and practice of Public Health surveillance, Oxford University Press, New York, 2000, pp. 176–193.

[17] Sharp S.J., Thompson S.G., Analysing the relationship between treatment effect and underlying risk in meta-analysis: comparison and development of approaches, Stat.Med. (2000) 19:3251–3274.

[18] Sweeting M.J., Sutton A.J., Lambert P.C., What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data, Stat. Med. (2004) 23:1351–1375.

[19] Thompson S.G., Smith T.C., Sharp S.J., Investigating underlying risk as a source of heterogeneity in meta-analysis, Stat. Med. (1997) 16:2741–2758.

[20] Tongue S.C., Wilesmith J.W., Nash J., Kossaibati M., Ryan J., The importance of the PrP genotype in active surveillance for ovine scrapie, Epidemiol. Infect. (2007) 25: 1–10.

[21] Warn D.E., Thompson S.G., Spiegelhalter D.J., Bayesian random effects meta-analysis of trials with binary outcomes: methods for the absolute risk difference and relative risk scales, Stat. Med. (2002) 21:1601–1623.