Prevalence of sheep infected with classical scrapie in Great Britain: integrating multiple sources of surveillance data for 2002

Simon Gubbins*

Pirbright Laboratory, Institute for Animal Health, Ash Road, Pirbright, Surrey GU24 0NF, UK

Estimates for the prevalence of sheep infected with classical scrapie are essential for assessing the efficacy of control strategies that have been implemented in Great Britain (GB). Here a back-calculation approach was used to estimate the prevalence in the GB national flock by integrating data on reported cases and the results of abattoir and fallen stock surveys for 2002. Prevalence estimates ranged from 0.33 to 2.06%, depending on the estimates used for the frequencies of prion protein (PrP) genotypes in the national flock and the stage of incubation at which the diagnostic tests used are able to detect infected animals. The risk of infection was found to be higher than that of clinical disease, especially in those PrP genotypes that have a later age at onset of clinical disease. Moreover, results suggest that a high proportion (more than 55%) of infected animals surviving to disease onset die on farm before clinical signs become apparent, which helps account for the high observed prevalence in the fallen stock compared with the abattoir survey. The analyses indicated that attention needs to be given to identifying the stage of incubation at which diagnostic tests are able to detect infected animals and obtaining better demographic data for the GB national flock.

Keywords: transmissible spongiform encephalopathy; scrapie; epidemiology; prion protein genotype; sheep; back calculation

1. INTRODUCTION

Scrapie is a fatal neurodegenerative disorder of sheep and goats and is a member of the transmissible spongiform encephalopathy (TSE) group of diseases, which also includes bovine spongiform encephalopathy (BSE) in cattle and variant Creutzfeldt–Jakob disease (vCJD) in humans. Wide-scale control measures for scrapie, including selective breeding programmes and action within affected flocks, have been introduced throughout the European Union in order to reduce the risk to human health posed by the possible presence of BSE in sheep (EC 2005). To monitor the effect of these control measures, it is essential to have estimates for the prevalence of scrapie.

There are several sources of data that allow the prevalence or incidence of scrapie in Great Britain (GB) to be estimated, but all have associated drawbacks. Statutory notification data provide one source for estimating the incidence of clinical disease (del Rio Vilas et al. 2006), but suffer from under-reporting (Hoinville et al. 2000; Sivan et al. 2003). Anonymous postal surveys conducted in 1998 and 2002 aimed to overcome the reluctance of farmers to report suspect cases, but the accuracy of the results depends on farmers’ ability to correctly recognize scrapie in their animals (Hoinville et al. 2000; Sivan et al. 2003). Alternatively, abattoir and fallen stock surveys can be used to estimate the prevalence of infection, but they rely on the detection of infected animals prior to the onset of clinical signs and, moreover, focus only on particular sections of the sheep population (Simmons et al. 2000; Webb et al. 2001; Gubbins et al. 2003; del Rio Vilas et al. 2005a; Elliott et al. 2005).

To overcome the problems associated with individual sources of surveillance data, it is possible to integrate different sources and, hence, produce more robust estimates. An integrative approach has recently been used to estimate the proportion of scrapie-affected flocks in GB (del Rio Vilas et al. 2005b), but corresponding methods have yet to be developed for the prevalence of scrapie at the animal level. By contrast, methods that integrate data from several sources have been developed for other TSEs, notably BSE in cattle and vCJD in humans. In the case of BSE, back-calculation methods were used to link data on the prevalence of infection in apparently healthy cattle and the incidence of confirmed clinical disease (Donnelly et al. 2002; Ferguson & Donnelly 2003), while for vCJD
similar methods were used to combine data on reported deaths and results from a survey of appendix tissues (Ghani et al. 2003; Clarke & Ghani 2005).

The aim of this paper was to estimate the prevalence of sheep infected with classical scrapie in GB by integrating data collected during 2002 on the incidence of reported clinical disease, the prevalence of infection in apparently healthy sheep slaughtered for human consumption and the prevalence of infection in sheep found dead on farm. However, scrapie has a strong host genetic component at the ovine prion protein (PrP) gene, which influences both the risk of infection and the incubation period (Detwiler & Baylis 2003; Baylis et al. 2004; Gubbins & Roden 2006; Tongue et al. 2006). Five alleles of the PrP gene (defined by the amino acids at codons 136, 154 and 171) are commonly found in British sheep, which, in the order of increasing risk of clinical disease and decreasing age at onset, are: ARR; AHQ; ARH; ARQ; and VRQ. Consequently, a back-calculation approach was adopted, which, although similar to those used previously for BSE and vCJD, also incorporates the effects of PrP genotype on the risk of infection and age at onset of disease (cf. Gubbins et al. 2003; Clarke & Ghani 2005).

The PrP genotypes were missing for a number of reported cases and negative samples from the FS and AS (75, 187 and 1121, respectively). The missing genotypes were inferred by assuming that they occurred at the same frequency as the known PrP genotypes in the same population.

2.2. Modelling approach

All animals were assumed to become infected at, or close to, birth because the risk of infection is the greatest during the perinatal period (Foster & Dickinson 1989; Hunter & Cairns 1998), and there is evidence for a decrease in susceptibility with age (St Rose et al. 2006). In this case, the probability that an animal of genotype $j$ develops clinical disease in age class $a$ (comprising animals between $a-1$ and $a$ years of age) is given by

$$c_{ja} = r_j\phi \int_{a-1}^{a} f_j(v) \, dv,$$

where $\phi$ is the baseline risk of infection; $r_j$ is the relative risk of infection for genotype $j$; and $f_j$ is the probability density function for the lognormal incubation period distribution (with genotype-specific parameters, $\mu_j$ and $\sigma_j$).

The incubation period for scrapie is long relative to the mean life expectancy of a sheep and, hence, it is essential to include survivorship when calculating the probability of an animal developing clinical disease. Moreover, it was assumed that a proportion of infected animals that survive to disease onset are sent to slaughter or die on farm before clinical signs become apparent (Donnelly et al. 2002; cf. Ferguson & Donnelly 2003). Finally, there is under-reporting of cases (Hoinville et al. 2000; Sivam et al. 2003) and, hence, it is necessary to consider the probability of a case being reported. Consequently, the expected number of reported cases in genotype $j$ and age class $a$ is given by

$$R_{ja} = \rho (1-K) s_a B_j r_j \phi \int_{a-1}^{a} f_j(v) \, dv,$$

where $\rho$ is the probability of reporting (assumed to be independent of age and PrP genotype); $K$ is the proportion of infected animals surviving to disease onset, which are sent to slaughter or die on farm before clinical signs become apparent; $s_a$ is the probability of surviving to be in age class $a$; and $B_j$ is the number of animals of genotype $j$ in a birth cohort.

The prevalence of infection in the FS or AS has two components: the first corresponds to infected animals found dead or sent to slaughter prior to the onset of clinical disease and the second to infected animals that survive to the onset of disease, but which die on farm or are sent to slaughter before clinical signs become apparent. The prevalence of infection in animals of

(i) The scrapie notifications database, held by the Veterinary Laboratories Agency (VLA), records details of all suspect and confirmed clinical cases reported in GB (del Rio Vilas et al. 2006). This was used to provide the age and PrP genotype of the number of animals found dead on farm (figure 1a,b). Full details of the FS have been presented elsewhere (del Rio Vilas et al. 2005a). Briefly, brainstem samples from 913 sheep over 18 months of age found dead on farm were screened by western blot or ELISA for the presence of PrPSc. Animals were sampled at 42 abattoirs with a throughput of at least 10 000 adult sheep per year; these abattoirs slaughtered 93% of adult sheep in GB.

(ii) Results of a fallen stock survey (FS) provided the number of animals found dead on farm sampled and the number of positive samples for each PrP genotype (figure 1c,d). Full details of the FS have been presented elsewhere (del Rio Vilas et al. 2005a). Briefly, brainstem samples from 913 sheep over 18 months of age found dead on farm were screened by western blot for the presence of disease-associated prion protein (PrPSc; taken to be an indicator of infectivity). Animal health offices throughout GB (but excluding the Shetland Isles) were allocated weekly quotas based on the sheep population in their catchment area.

(iii) Results of an abattoir survey (AS) gave the number of animals sampled and the number of positive samples for each PrP genotype (figure 1e,f). Full details of the AS have been presented elsewhere (Elliott et al. 2005). Briefly, brainstem samples from 30 115 apparently healthy sheep over 18 months of age were

J. R. Soc. Interface (2008)
genotype $j$ and age class $a$ found dead on farm is

$$p_{ja}^{(FS)} = r_j \phi \left\{ \eta_a (s_a - s_{a+1}) \int_{a-1}^{a} \int_w f_j(w) \, dw \, dv \right\}
+ \bar{\xi} \eta_a s_a K \int_{a-1}^{a} f_j(v) \, dv,$$

(2.3)

while the corresponding prevalence for animals sent to slaughter is

$$p_{ja}^{(AS)} = r_j \phi \left\{ (1 - \eta_a) (s_a - s_{a+1}) \int_{a-1}^{a} \int_w f_j(w) \, dw \, dv \right\}
+ (1 - \bar{\xi} \eta_a) s_a K \int_{a-1}^{a} f_j(v) \, dv,$$

(2.4)
where \( w \) is the age at onset of clinical disease; \( \eta_a \) is the proportion of uninfected animals that are found dead on farm in age class \( a \); and \( \xi \) is the relative risk of an infected animal being found dead on farm. However, the diagnostic tests used in the surveys are not perfect and, hence, not every infected animal will be detected. If it is assumed that test sensitivity depends on the time to the onset of clinical disease, the detectable prevalence for genotype \( j \) and age class \( a \) in animals found dead on farm is given by

\[
d^{(FS)}_{ja} = r_j \phi \left[ \eta_j (s_a - s_{a+1}) \int_{a-1}^{\infty} \xi(v, w) f_j(w) \, dw \, dv \right.
+ \xi \eta_j s_a K \int_{a-1}^{\infty} \xi(v, w) f_j(w) \, dv \right],
\]

while for animals sent to slaughter, it is

\[
d^{(AS)}_{ja} = r_j \phi \left[ (1 - \eta_j) (s_a - s_{a+1}) \int_{a-1}^{\infty} \xi(v, w) f_j(w) \, dw \, dv \right.
+ (1 - \xi \eta_j) s_a K \int_{a-1}^{\infty} \xi(v, w) f_j(w) \, dv \right],
\]

where \( \xi(v, w) \) is the probability of detecting an infected animal tested at age \( v \) given that it would have developed clinical disease at age \( w \). It was assumed that an infected animal would be detected, provided it was in the final proportion of the incubation period, so that the probability of detection, \( \xi(v, w) \), is given by

\[
\xi(v, w) = \begin{cases} 0 & v \leq (1 - \delta) w, \\ 1 & v > (1 - \delta) w, \end{cases}
\]

where \( \delta \) is the preclinical detection proportion (cf. Webb et al. 2001; Donnelly et al. 2002; Gubbins et al. 2003; Clarke & Ghani 2005). Finally, the proportion of uninfected animals of genotype \( j \) in age class \( a \) found dead on farm is

\[
u^{(FS)}_{ja} = (1 - r_j \phi) \eta_j (s_a - s_{a+1}),
\]

while that for animals sent to slaughter is

\[
u^{(AS)}_{ja} = (1 - r_j \phi) (1 - \eta_j) (s_a - s_{a+1}).
\]

The prevalence of infection in the GB national flock, \( p^{(op)} \), is found by computing the prevalence in each age and genotype class multiplied by the proportion of the national flock in that class, so that

\[
p^{(op)} = \sum_j \sum_a \pi_{ja} B_j,
\]

where

\[
\pi_{ja} = s_a r_j \phi \int_{a-1}^{\infty} f_j(w) \, dw.
\]

is the prevalence of infected sheep of genotype \( j \) in age class \( a \).

2.3. Maximum-likelihood methods

Scrapie notifications data provide the number of reported cases \( X_{ja} \) of genotype \( j \) in age class \( a \) (figure 1a,b). The observed number of reported cases was assumed to follow a Poisson distribution with the mean given by the expected number of reported cases \( (R_{ja} \) defined by equation (2.2)),). Because the age-at-onset parameters were estimated independently of the 2002 data (see below), only the total number of cases of each PrP genotype were used for estimation, in which case the log likelihood \( l^{(op)} \) is

\[
l^{(op)} = \sum_j \left[ -\sum_a R_{ja} + \sum_a X_{ja} \log \left( \sum_a R_{ja} \right) \right]

\]

The prevalence of infection in the GB national flock, \( p_{op} = \sum_j \sum_a \pi_{ja} B_j \), is the last parameter to be estimated, along with the age at onset of classes \( a \) in each age class; and \( B_a \) number of sheep of genotype \( j \) in a birth cohort) were estimated using data derived from a number of sources (see the electronic supplementary material, appendix A). The age-at-onset parameters \( \mu_a \) and \( \sigma \) were estimated from data on the age at onset of cases reported between July 1998 and December 2005 (see the electronic supplementary material, appendix B). The preclinical detection proportion \( \delta \) was estimated using the data from
Table 1. Estimates for the prevalence of sheep infected with classical scrapie in GB and the proportion of infected animals surviving to disease onset, which are sent to slaughter or die on farm before clinical signs become apparent, and goodness-of-fit statistics for four scenarios depending on the estimates used for the frequency of PrP genotypes in the GB national flock and the preclinical detection period ($\delta$).

<table>
<thead>
<tr>
<th></th>
<th>IAH frequencies</th>
<th></th>
<th>NSP frequencies</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\delta=25%$</td>
<td>$\delta=50%$</td>
<td>$\delta=25%$</td>
<td>$\delta=50%$</td>
</tr>
<tr>
<td>population prevalence (% sheep infected; $p_{\text{POP}}$)</td>
<td>2.06</td>
<td>0.67</td>
<td>0.98</td>
<td>0.33</td>
</tr>
<tr>
<td>maximum-likelihood estimate (MLE)</td>
<td>(0.90, 4.66)</td>
<td>(0.30, 1.49)</td>
<td>(0.41, 2.35)</td>
<td>(0.15, 0.75)</td>
</tr>
<tr>
<td>proportion (%) of infected animals surviving to disease onset, which die on farm before clinical signs become apparent ($K$)</td>
<td>91.98</td>
<td>76.71</td>
<td>83.73</td>
<td>55.07</td>
</tr>
<tr>
<td>MLE 95% CI</td>
<td>(88.78, 94.11)</td>
<td>(67.44, 82.95)</td>
<td>(77.41, 87.98)</td>
<td>(38.54, 66.68)</td>
</tr>
<tr>
<td>goodness-of-fit measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Akaike information criterion (AIC)</td>
<td>125.59</td>
<td>133.24</td>
<td>112.97</td>
<td>121.39</td>
</tr>
<tr>
<td>$\chi^2$ statistic$^a$</td>
<td>20.14</td>
<td>18.16</td>
<td>15.60</td>
<td>14.39</td>
</tr>
<tr>
<td>degrees of freedom (d.f.$^a$)</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>$P$ value$^a$</td>
<td>0.13</td>
<td>0.20</td>
<td>0.34</td>
<td>0.42</td>
</tr>
</tbody>
</table>

$^a$ $\chi^2$ goodness-of-fit statistics are shown only for reported cases and AS positives owing to the small number of observed and expected FS positives in all PrP genotypes (figure 2).

pathogenesis experiments, while the specificity of the diagnostic test ($\psi$) was estimated to be 100% (see the electronic supplementary material, appendix C). Finally, the probability of reporting a case ($\rho$) was obtained from the results of an anonymous postal survey conducted in 2002, which suggested that 38% of farmers who had suspect scrapie cases reported them (Sivam et al. 2003; see also Böhning & del Rio Vilas 2008). Thus, a total of 17 parameters were estimated directly from the 2002 surveillance data.

2.5. Sensitivity analysis

No unbiased population-level PrP genotype data are available for the GB national flock (Tongue et al. 2006). Two sources of data were used to estimate the number of sheep of each PrP genotype in a birth cohort: sheep sampled as part of the National Scrapie Plan for GB (NSP; Eglin et al. 2005) and sheep from scrapie-affected and unaffected flocks sampled as part of a case-control study run by the Institute for Animal Health (IAH: Baylis et al. 2004; see the electronic supplementary material, appendix A). Limited pathogenesis data suggest that the preclinical detection proportion ($\delta$) is likely to lie somewhere between 25 and 50% (see the electronic supplementary material, appendix C). To assess the sensitivity of the parameter estimates to these sources of uncertainty, the maximum-likelihood methods were implemented for four scenarios that differed in the population-level PrP genotype data (IAH or NSP) and the preclinical detection proportion ($\delta=25$ or 50%).

3. RESULTS

Estimates for the prevalence of sheep infected with classical scrapie in GB ranged from 0.33 to 2.06% (table 1). A higher prevalence estimate and wider confidence interval (CI) were obtained if a shorter preclinical detection proportion was assumed. Similarly, a higher estimate was obtained if the IAH genotype frequencies were used instead of the NSP frequencies. The estimates for the relative risk of infection in each PrP genotype depended on the preclinical detection proportion, though the impact of the different population-level PrP genotype frequencies was much greater, with higher estimates associated with the IAH frequencies (figure 2; estimates are given in full in the electronic supplementary material, appendix D). However, the same PrP genotypes were associated with the highest risk regardless of the PrP genotype frequencies used (figure 2). For all scenarios, the estimated relative risk of an infected animal being found dead on farm was such that virtually all cases that die before the onset of overt clinical signs would be found dead on farm rather than sent to slaughter (i.e. $\xi_{PA}=1$). Moreover, the results suggest that the proportion of infected animals surviving to disease onset which die on farm before clinical signs become apparent is high (more than 55%; table 1).

There was good agreement between the observed and expected number of reported cases in all PrP genotypes (figure 3a). Similarly, the expected number of positive samples in the AS was similar to that observed, except in ARQ/ARQ for which the number of positives is consistently overestimated (figure 3c). Formal $\chi^2$ goodness-of-fit tests indicated no significant ($P>0.05$) differences between observed and expected numbers of reported cases and positive samples in the AS (table 1). However, the number of positive samples in the FS was consistently underestimated for those PrP genotypes for which positive samples were detected (figure 3b).

4. DISCUSSION

Three previous estimates for the prevalence of sheep infected with classical scrapie in GB have been derived, all using AS data: 0.22% (95% CI: 0.01–0.97%) based on a survey conducted in 1997/1998 (Gubbins et al. 2003);
resulted in underestimation of the prevalence, given incomplete survival. Both these omissions will have incorporation of PrP genotype nor did the estimates used for narrower (cf. table 1). None of these analyses incorporated PrP genotype nor did the estimates used for the age-at-onset parameters allow for the effect of incomplete survival. Both these omissions have resulted in underestimation of the prevalence, given the effects of PrP genotype and survivorship on the risk of infection and the age at onset of clinical disease identified in the present study. Their omission is also likely to have resulted in a spurious precision to the prevalence estimates.

Incorporating PrP genotype into the analyses allowed the risk of infection in individual genotypes to be estimated. The risk was the highest in VRQ/VRQ (figure 2), which is also associated with the highest risk of clinical disease (Baylis et al. 2004; Tongue et al. 2006). The ranking of genotypes by risk of infection was similar to the ranking by risk of clinical disease, except for ARR/VRQ where the risk of infection was much higher (figure 2). With the exception of the ARH/VRQ genotype, however, the relative risk of infection was typically higher than that of clinical disease (figure 2; cf. Baylis et al. 2004; Gubbins & Roden 2006; Tongue et al. 2006). Moreover, the greatest discrepancies between the relative risk of infection and that of disease occurred in those PrP genotypes that have a later age at onset (figure 2; see the electronic supplementary material, appendix B), reflecting the fact that relatively few infected animals of these PrP genotypes survive to disease onset.

Differences in the estimates obtained for the relative risk of infection when using different population-level PrP genotype frequencies (IAH or NSP; figure 2) can be explained by biases in these datasets (e.g. Tongue et al. 2006). The IAH dataset was derived from approximately 30 affected and 30 unaffected flocks (Baylis et al. 2004; Goldmann et al. 2005). Because affected flocks tend to have different PrP genotype profiles from unaffected flocks and, in particular, higher frequencies of PrP genotypes associated with a high risk of clinical disease (Baylis et al. 2000; Tongue et al. 2004; Goldmann et al. 2005), the IAH dataset is likely to overestimate the frequency of PrP genotypes associated with a higher risk of scrapie and underestimate the frequency of those associated with a lower risk. Conversely, the NSP dataset was derived from a voluntary ram genotyping scheme in pure-bred flocks. This scheme requires culling or castration of VRQ-bearing rams and, hence, may underestimate the frequency of these PrP genotypes. Moreover, a large number of Texel rams have been genotyped as part of the NSP and this breed has a very high frequency of the ARH allele (Eglin et al. 2005). Consequently, the frequency of this allele in the national flock is likely to be overestimated by the NSP dataset.

The robustness of the prevalence estimates depends on the estimates for the sensitivity of the diagnostic tests used in the AS and FS and the probability of reporting clinical disease. Although estimates are available for the sensitivity of diagnostic tests when used on confirmed clinical cases (Philipp et al. 2005), it is more difficult to determine the sensitivity of the tests in infected, but preclinical animals. In this study, two values were used for the proportion of the incubation period during which the test is able to detect infected animals (β; 25 and 50%), reflecting the limited data available from pathogenesis studies. Comparison of the AIC for the best-fit models
provides some evidence that $d = 25\%$ results in a better model fit than $d = 50\%$ (table 1). However, an independent estimate for this critical parameter is essential when interpreting the surveillance data.

The probability of reporting is an essential parameter if the model is to fit data both on notified cases and from the AS, because the abattoir data yield a prevalence estimate that is much higher than would be inferred from reported clinical disease alone. This issue has also been identified when analysing surveillance data for BSE in cattle (Donnelly et al. 2002; Ferguson & Donnelly 2003) and for vCJD in humans (Ghani et al. 2003; Clarke & Ghani 2005). However, the results of the present analysis suggest that there is likely to be underascertainment of scrapie cases for reasons other than under-reporting. In particular, the model predicts that a high proportion (more than 55\%) of infected animals surviving to disease onset die on farm before clinical signs become apparent (table 1). A similarly high level of underascertainment was identified for BSE in cattle, which was also related to infected animals being sent to slaughter or dying on farm before disease onset (Donnelly et al. 2002; Ferguson & Donnelly 2003). There was also an increased risk of being found dead for BSE-infected cattle (Ferguson & Donnelly 2003), though the magnitude of the increase was lower than that for scrapie-infected sheep.

The high proportion of infected animals surviving to disease onset which die on farm before clinical signs become apparent accounts for why the FS has a markedly higher prevalence of infection than the AS (figure 1; cf. del Rio Vilas et al. 2005a; Elliott et al. 2005). This finding supports the contention that being found dead should be considered a sign of scrapie (Clark & Moar 1992; Clark et al. 1994). Furthermore, it helps explain the observation that scrapie-affected flocks have a higher frequency of animals that are found dead than unaffected flocks (McLean et al. 1999).

The aim of this paper was to integrate multiple sources of scrapie surveillance data, which was achieved using a back-calculation approach. Combining different sources of data helps to correct for biases associated with individual sources and, hence,
produces more robust prevalence estimates. Moreover, the modelling approach developed in the present study includes the effect of PrP genotype on the risk of infection and the age at onset of disease, which is essential if it is to be extended to examine temporal trends in prevalence. In particular, any changes in the PrP genotype profile of the GB national flock (e.g. as a result of control measures) must be taken into account when interpreting surveillance data from multiple years.

The analysis has also served to highlight areas of uncertainty and their impact on the interpretation of surveillance data. This has shown that greater attention needs to be given to identifying the stage of incubation at which diagnostic tests are able to detect infected animals, and obtaining better demographic data for the GB sheep population and, in particular, the frequency of PrP genotypes in the national flock.

This work was funded by the Department for Environment, Food and Rural Affairs (Defra). The author is grateful to Victor del Rio Vilas and Julia Colvin (both Veterinary Laboratories Agency; VLA) for providing scrapie surveillance data and to Charlotte Cook and Rachel Eglin (both VLA) for providing data on the frequency of animals found dead on farm.

**REFERENCES**


St Rose, S. G. et al. 2006 Comparative evidence for a link between Peyer’s patch development and susceptibility to transmissible spongiform encephalopathies. BMC Infect. Dis. 6, 5. (doi:10.1186/1471-2334-6-5)

