Updated projections of future vCJD deaths in the UK
Azra C Ghani*, Christl A Donnelly, Neil M Ferguson and Roy M Anderson

Address: Department of Infectious Disease Epidemiology, Faculty of Medicine, Imperial College of Science, Technology and Medicine, Norfolk Place, London W2 1PG, UK
Email: Azra C Ghani* - a.ghani@imperial.ac.uk; Christl A Donnelly - c.donnelly@imperial.ac.uk; Neil M Ferguson - neil.ferguson@imperial.ac.uk; Roy M Anderson - roy.anderson@imperial.ac.uk
* Corresponding author

Abstract

Background: Past projections of the future course of the vCJD epidemic in the UK have shown considerable uncertainty, with wide confidence bounds. However, recent vCJD case data have indicated a decrease in the annual incidence of deaths over the past two years.

Methods: A detailed survival model is fitted to the 121 vCJD deaths reported by the end of 2002 stratified by age and calendar time to obtain projections of future incidence. The model is additionally fitted to recent results from a survey of appendix tissues.

Results: Our results show a substantial decrease in the uncertainty of the future course of the primary epidemic in the susceptible genotype (MM-homozygous at codon 129 of the prion protein gene), with a best estimate of 40 future deaths (95% prediction interval 9–540) based on fitting to the vCJD case data alone. Additional fitting of the appendix data increases these estimates (best estimate 100, 95% prediction interval 10–2,600) but remains lower than previous projections.

Conclusions: The primary vCJD epidemic in the known susceptible genotype in the UK appears to be in decline.

Background

Since the announcement in March 1996 of a link between variant Creutzfeldt-Jakob disease (vCJD) and bovine spongiform encephalopathy (BSE) in cattle [1–3], there has been considerable concern regarding the potential for an epidemic of vCJD in the UK. To date, there have been 129 cases of vCJD reported in the UK, 7 of which remain alive, with further cases reported in Ireland, France, Italy and Canada. The recent decrease in deaths from vCJD in the UK (17 cases in 2002, compared with 20 in 2001 and 28 in 2000) has prompted suggestions that any epidemic of vCJD may indeed be much smaller than previously feared. However, a recent survey of appendix tissues surgically removed between 1995 and 1999 showed a relatively high prevalence of infection (1 positive sample in 8318 tissues) [4]. The latter result has raised concern over the theoretical possibility of secondary transmission of vCJD through surgical instruments or blood products [5,6].

In this paper we use a previously published model [7–12] which relates estimated exposure to BSE-infected animals to the vCJD case data to provide updated estimates and prediction intervals for the short- and long-term course of the primary vCJD epidemic in the UK, based on fitting to vCJD deaths occurring to the end of 2002 and additionally to the preliminary results from the survey of appendix tissues.
Methods

**vCJD case data**

Our updated analyses are based on the vCJD deaths occurring to the end of December 2002. 121 deaths due to vCJD had been diagnosed in the UK at this time (Figure 1) http://www.doh.gov.uk/cjd/. The median age at onset of symptoms is 26 years, and at death 28 years. The excess of cases in younger individuals, compared with what would have been expected in a random sample of the population, suggests that younger individuals are either more susceptible to infection, were exposed to a greater extent, and/or have shorter incubation periods.

All cases tested to date are methionine (MM) homozygous at codon 129 of the prion protein (PrP) gene [13]. Approximately 40% of the Caucasian population share this genetic trait, with 13% valine homozygous and the remaining 47% heterozygous [14,15]. This genetic marker is known to be important for other human Transmissible Spongiform Encephalopathies (TSEs) including sporadic CJD, iatrogenic CJD and kuru [16–18]. Whilst it is possible that other genotypes are also susceptible to vCJD (and are either less susceptible than the MM genotype, or have longer incubation periods), at present we have no cases with which to constrain potential epidemic scenarios among these other genotypes. Our model predictions are therefore limited to the 40% of the population with the MM genotype.

**vCJD prevalence data**

Preliminary results from testing appendix tissues for the presence of abnormal PrP showed 1 positive appendix in a sample of 8318 tissues [4]. The study was interrupted with this result reported due to public health interest. The majority of tissues tested (70%) were from the 20–29 age group (with the remaining tissues restricted to the 10–50 age-group), and all were removed between 1995 and 1999. The data were analysed in batches of approximately 1000 samples to ensure anonymity. These data provide the first estimate of the prevalence of asymptomatic infection in the population in the highest risk age-group. The observation translates to an estimated detectable prevalence of 120 per million with 95% confidence interval 0.5–892 per million across the whole sample [4]. However, interpretation of this result remains uncertain since it is not known how sensitive the tests are over the course of the incubation period, and whether a positive test is indicative of an infected individual who will go on to develop clinical disease. Whilst evidence of abnormal PrP has been found in the appendix tissue of 19 out of 20 symptomatic vCJD cases examined to date, as well as in 2 cases with appendices removed prior to the onset of symptoms (3 and 4 years before death respectively), it was not found in the appendix tissue of one patient removed 9 years prior to the onset of symptoms [4,19]. We assume in our analyses that the positive appendix arose from an MM-homozygous individual. If this is not the case, then projections for the MM-homozygous population are similar to those obtained fitting to the case data alone (data not shown). This would also indicate the potential for vCJD cases to arise in other genotypes. However, since there have not been any vCJD cases in these other genotypes to date, it is difficult at present to project the potential scale of any epidemic in these populations without making additional untestable assumptions regarding these other genotypes.

**Statistical model**

Full details of the model and fitting procedure, and detailed sensitivity analyses, have previously been published [7–12]. In brief, the probability that an individual dies from clinical disease at time $u$ and age $a$ is given by

$$p(u, a) = S(u, a) \int_{u=a}^{u} f(u-t) \beta(t, a-u) \exp \left[ -\int_0^{t} \beta(t', a-u+t') dt' \right] dt$$ (1)

where $S(u,a)$ is the survival probability (estimated from UK census data), $f(u)$ is the incubation period distribution and $\beta$ is the transmission coefficient. The time- and age-dependent exposure hazard $I(t, a)$ is given by

$$I(t, a) = v(t) g(a) \int \Omega(z) w(z, t) dz.$$ (2)
Here \( \nu(t) \) is the effectiveness of control measures limiting the bovine tissues allowed into the human food supply (introduced mid-1989, and allowed to vary between 0 and 100% effective at reducing infectivity from this point onwards), \( g(a) \) an age-dependent susceptibility/exposure function, \( \Omega(z) \) the relative infectiousness of bovines over the course of their incubation period and \( w(z, t) \) the proportion of cattle slaughtered at time \( t \) and time \( z \) away from disease onset. The expected number of cases at time \( u \) and age \( a \) is given by

\[
c(u, a) = B(u - a) \rho(u, a) \tag{3}
\]

where \( B(u - a) \) is the number of individuals born at time \( u \) - \( a \).

The model is fitted using maximum likelihood methods to the time- and age-stratified vCJD deaths to the end of 2002 assuming that the data arise from a Poisson distribution. The individual cases are categorised into cells with at least 5 observations in each cell. The best-fit point is that which maximises the log likelihood (LnL), or equivalently minimises

\[
-2\text{LnL} = 2 \sum_{u,a} c(u,a) - x(u,a) \ln(c(u,a)) - \ln(x(u,a)!) \tag{4}
\]

where \( x(u,a) \) are the observed cases at time \( u \) and age \( a \). For any model parameter, maximum likelihood confidence intervals are obtained using likelihood ratio tests. Non-linear optimisation techniques are used to fit the model [20] using custom-written code. An intensive search of parameter space was performed, fitting from multiple starting points and restricting parameter bounds, to ensure that the best-fitting models were obtained.

Confidence intervals for the expected number of cases are obtained by re-parameterising the model so that the expected number of cases between times \( u_1 \) and \( u_2 \), \( T(u_1, u_2) \), is a model parameter. Combining equations 1–3, this is given by:

\[
T(u_1, u_2) = \int_{u_1}^{u_2} \left[ \int \left( B(v-u) \rho(v, u) \right) f(v-u) \right] \left( e^{\Omega(z) w(z, v-u)} \right) \mid_{z=0}^{z=\infty} \frac{1}{\beta} \left( 1-e^{-\beta x} \right) \text{d}x \tag{5}
\]

For each proposed parameter set, equation (5) was solved numerically to calculate \( \beta \). If no solution was possible (for example, if the proposed parameter set generated fewer infections than the required epidemic size) then the proposed parameter set was rejected. Prediction intervals are obtained by adding Poisson variability to the confidence intervals on the mean.

The model is additionally fitted to the prevalence data obtained from a recent survey of appendix tissues. As the study was interrupted (although not stopped) due to the observation of a positive appendix, the 95% confidence interval for the prevalence of infection in the population is obtained from the likelihood term for a negative binomial distribution. The reported results arise from 9 batches with the positive result arising in the sixth batch. The likelihood for this result (assuming a binomial distribution for the sixth batch conditional on having observed a positive result and for the final three batches) is therefore given by

\[
\prod_{i=1}^{5} (1-p_i)^{n_i} \prod_{i=6}^{9} \binom{n_i}{x_i} \left( \frac{n_i!}{x_i! (n_i-x_i)!} \right) p_i^{x_i} (1-p_i)^{n_i-x_i} \tag{6}
\]

where \( p_i \) is the expected prevalence in batch \( i \), \( n_i \) is the number of samples in batch \( i \) and \( x_i \) is the number of positive appendixes in batch \( i \). The expected prevalence at time \( u \) in age-group \( a \) is given by

\[
\prod_{u=0}^{u} \int_B(u-a)S(u, a)(v-u)B(v, a-u+v)\exp \left[ -\int_0^z \beta(t', a-u+v)dt' \right] \text{d}u \tag{7}
\]

The expected prevalence in batch \( i \) is calculated from this expression as the average prevalence in the age-group over the period of time from which the samples were collected. The model fit is then judged by the combination of the

### Table 1: Projected future number of vCJD cases (MLE and 95% prediction interval) based on fitting to the vCJD case data alone and to the case data and prevalence data simultaneously.

<table>
<thead>
<tr>
<th></th>
<th>vCJD case data alone</th>
<th>vCJD case data and prevalence data simultaneously</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003–2005</td>
<td>20 (3–70)</td>
<td>28 (4–72)</td>
</tr>
<tr>
<td>2003–2008</td>
<td>40 (8–138)</td>
<td>60 (9–152)</td>
</tr>
<tr>
<td>2003–2080</td>
<td>40 (9–540)</td>
<td>100 (10–2,600)</td>
</tr>
<tr>
<td>Fit of the model (-2LnL)</td>
<td>18.7</td>
<td>26.6</td>
</tr>
<tr>
<td>Goodness-of-fit p-value</td>
<td>0.29</td>
<td>0.32</td>
</tr>
</tbody>
</table>
Poisson likelihood fit to the cases and the likelihood fit to the prevalence data given by equation (6).

Both the sensitivity and specificity of the diagnostic tests at different incubation stages are unknown. We include an additional parameter which specifies the proportion of the incubation period for which the test is assumed to be fully sensitive (working backwards from death), and assume that the test has no sensitivity prior to this time. We also assume the test is fully specific. Throughout this work we assume that the positive appendix is from an MM-homozygous individual, but that the 8318 tissues are representative of the general population (and hence 40% of these tissues are from MM-homozygous individuals).

**Parameters**

Predictions from the model are sensitive to the estimates of exposure to BSE-infected material, $w(z,t)$ [9]. As in previous work, our analyses rely on estimates of exposure using data relating to BSE in cattle in Great Britain. However, in our most recent analyses estimating exposure to preclinical infected animals and unreported clinical cases we have integrated the analysis of BSE clinical case data (up to October 2002), screening data from apparently healthy cattle and screening data from high risk’ (casualty/fallen stock) cattle. Previous work [21] details the methodological extensions required to integrate the analysis of data on clinical case incidence and results from the screening of apparently healthy cattle to obtain estimates of case ascertainment rates. Key to this is the incorporation of a mechanism underlying the apparent underascertainment of cases indicated by the screening data. Two different possible mechanisms were examined: differential mortality of BSE-infected animals and underreporting of clinical cases [21]. The former is better able to explain the observed data and involves animals close to clinical onset being more likely than their age-matched peers to die on farm or be slaughtered. Thus, differential mortality leads to a higher infection prevalence in slaughtered animals than would be expected naively from the clinical case incidence data alone. Furthermore, animals slaughtered as a result of the excess risk imposed by differential mortality, all in the latest stages of incubating the disease, are allowed via a fitted parameter to be more likely to be categorized as a high-risk animal (either as fallen stock found dead on farms or casualty animals). If the parameter is significant, then the resulting prevalence of infection is greater in high-risk animals than in other animals slaughtered as apparently healthy.

The exposure estimates used here were obtained assuming that diagnostic test sensitivity was 100% for the last 3 months of the incubation period but virtually 0% before this and assuming that deaths due to differential mortality were distributed over the 3 months of the incubation period of infected animals (this is the differential mortality pattern assumed in Donnelly et al. 2002 [21]). Furthermore, the diagnostic test was assumed to be 100% specific in all animals.

Pathogenesis experiments in cattle have indicated that BSE infectivity is most widely distributed in the carcass towards the end of the incubation period [22,23]. We therefore model the relative infectiousness of cattle at different stages of incubation assuming that infection increases exponentially towards the end of the incubation period [7]. Clinical BSE cases (including those reported prior to mid-1988) are assumed to be as infectious as those pre-clinical animals within 6 months of onset of disease. Predictions are fairly insensitive to the parameters for this parametric form.

In previous work, we explored a number of different parametric forms for the age-dependent susceptibility/exposure function, $g(a)$ [9]. The best fitting distributional form is a 3-parameter combined uniform distribution with gamma-distributed tails, which in its limits is either the uniform or gamma distribution.

The functional form of the incubation period distribution is informed by patterns observed for other TSEs. In particular, for other human TSEs it is known that the incubation period is long and highly variable [16,24,25]. For many TSEs there is a minimum incubation period so that functional forms must allow for the potential for a substantial delay following infection until the onset of clinical symptoms of vCJD [26,27]. We use a modified form of the four-parameter Generalised Lambda distribution [28] (with inverse CDF $F(p) = \frac{\lambda_1}{\lambda_2}(p^{\lambda_3} + (1 - p)^{-\lambda_3} - 1)/\lambda_2$) with inverse $\lambda_2$ CDF $g(p) = \frac{\lambda_1}{\lambda_2}(p^{\lambda_1} + (1 - p)^{-\lambda_1} - 1)/\lambda_1$. This distribution is more flexible than distributions based on standard functional forms (such as the Weibull, LogNormal or Gamma distributions), and incorporates all of these standard distributions as special cases.

**Results**

The revised projections for future numbers of cases in the susceptible genotype arising from a primary epidemic (i.e. from consumption of beef and beef products from infected animals) are shown in Table 1. Based on fitting to the vCJD case data alone (which is comparable to previous analyses [9,10]), the upper limit of the 95% prediction interval (PI) for future cases has decreased from approximately 7,000 cases last year to just 540 cases. Furthermore, the best estimate of future case numbers is now relatively small (40 future cases), indicating that the vCJD epidemic in this genotype appears to have peaked and is now in decline. This decline is also reflected in the short-term projections, with only 20 cases expected in the next 2 years (95% PI 3–70) and 40 cases in the next 5 years (95% PI 8–138).
If the model is additionally fitted to the prevalence data obtained from the survey of appendix tissues, and it is assumed that the positive appendix arose from an MM-homozygous individual who will go on to develop clinical disease, projections of the magnitude of the primary epidemic are increased. In particular, the best estimate of future cases is roughly doubled (from 40 to 100) whilst the upper 95% prediction bound is substantially increased from 540 cases to 2,600 cases (Table 1). Figure 2 shows the likelihood profiles for the fit of the model to the vCJD case and appendix data simultaneously and demonstrates the inconsistency between the vCJD case data and the prevalence data. The best fits to the vCJD case data are obtained for small epidemic sizes. In contrast, much better fits to the appendix data are obtained for larger epidemic sizes. The simultaneous fit of the model to both data sources is therefore a trade-off.

Addition of the appendix data to the model also affects the estimated incubation period distribution (Table 2). Figure 3 shows the maximum likelihood estimate (MLE) of the incubation period distribution obtained by fitting to the vCJD case data alone, and to the vCJD case and prevalence data simultaneously. The addition of the prevalence data to the model increases the estimated mean incubation period and results in a longer tail to the distribution. In contrast, as the prevalence data are obtained from the highest-risk age groups (10 to 30 years), the addition of these data to the model does not affect the MLE of age-dependent susceptibility/exposure (Figure 4), which remains similar to that obtained in previous work [9]. Our results continue to highlight that the young are either more susceptible or were exposed to a greater extent. Recent studies of the consumption patterns of high-risk beef products (including head meat and burgers) [29,30] have been unable to explain this solely in terms of exposure, suggesting that younger individuals may have also been more susceptible to infection.

Figure 5 shows the likelihood profile for the mean incubation period. In contrast to analyses undertaken last year [9], the data now allow a meaningful estimate of the mean incubation period to be made. By fitting to the case data alone, we estimate a mean incubation period of 12.6 years (95% confidence interval 8–35 years) for the MM-genotype. Additionally fitting the model to the prevalence data increases this mean to 16.7 years (95% confidence interval 8–30 years). These estimates are greater than those obtained for iatrogenic CJD (mean incubation period of 9–10 years with an upper 95% confidence limit of 15–16 years [24]) as would be expected from animal experiments of other TSEs in which longer incubation periods are obtained for infection via the oral route [31]. However, they are similar to recent estimates of the incubation period for kuru (mean 10–13 years, 90th percentile between 21 and 27 years [25]).

Discussion
The continued decrease in the annual incidence of vCJD in the UK has resulted in revised projections which are considerably lower than in previous years. In particular, the uncertainty in the potential size of the primary epi-
The vCJD epidemic in the known susceptible genotype (MM-homozygotes) has decreased resulting in narrower prediction intervals. Our results suggest that the vCJD epidemic will continue to decline with a best estimate of only 40 future cases and upper limit of the 95% prediction interval of 500 future cases. Even with the additional fitting of the appendix data (which estimates a higher prevalence of asymptomatic infection in the population than would be predicted by fitting the model to the vCJD case data alone), the best fitting model estimates just 100 future cases with an upper limit of the 95% prediction interval of 2,500 cases. These results show a substantial decline in the worst-case epidemic even compared to one year ago [9], and indicate that the epidemic now appears to be in decline. The narrower prediction intervals obtained, and in particular the reduction in the upper 95% prediction interval (from approximately 50,000 based on cases to the end of 2000 and 7,000 based on cases to the end of 2001) are due solely to the decline in vCJD cases observed over the past 2 years and to the stable age distribution of these cases.

Our analyses relate only to the 40% of the population known to be genetically susceptible to infection (MM-homozygotes). To date, there have been no cases of vCJD in other genotypes and hence it is difficult to make projec-

**Figure 3**
MLE of the probability density function (PDF) of the incubation period distribution obtained from fitting to the vCJD case data alone and to the vCJD case and prevalence data simultaneously.

**Figure 4**
MLE of the probability density function (PDF) of the age-dependent susceptibility/exposure function obtained from fitting to the vCJD case data alone and to the vCJD case and prevalence data simultaneously.

**Figure 5**
Likelihood profile for the mean incubation period. The bold lines show the cut-offs corresponding to the 95% confidence limits (based on the maximized log likelihood and one degree of freedom).
tions of the scale of any epidemic in the remainder of the population. Throughout our analyses we have assumed that the one positive appendix arose from an MM-homozygous individual. If this is not the case, then the results from the appendix study could indicate the presence of infection in other genotypes. Whilst it remains possible that cases may arise in these other genotypes in the future, it is unlikely that large epidemics would occur. This is because these genotypes must either have longer incubation periods or be less susceptible to infection to have resulted in no observed cases to date. In the former case, longer incubation periods are likely to result in a lower per-capita incidence because of the increased probability as they age that these individuals will die from other causes. In the latter case, reduced susceptibility will result in a lower per-capita incidence. The only way in which a large future epidemic could therefore occur is if the vCJD cases diagnosed to date are from a very small group of highly susceptible individuals (defined by a rare, and currently unknown, genetic marker) and the remainder of the population are less susceptible but will eventually develop disease. Such a scenario appears unlikely.

The analyses presented here do not provide any indication of the potential for secondary transmission of the agent, for example through blood products or via surgical instruments. Preliminary data from the study of appendix tissue suggest that the prevalence of asymptomatic infection may well be significant, although one must be cautious in extrapolating from small samples. Population-based studies of tonsil tissues are now planned (testing approximately 50,000 tissues) and a national archive is in preparation [32]. Although the results of these studies will provide useful information on the prevalence of late-stage infection, their interpretation will remain limited by our relatively poor understanding of the sensitivity and specificity of the diagnostic tests throughout disease incubation. Furthermore, such studies are necessarily limited in scale by both the samples available and the limited automation of the detection process. The development of a diagnostic test (for example, a blood test) that is able to detect infection early in the incubation period and that can be used to reduce the risk of secondary transmission therefore remains a high priority.

Competing interests
None declared.

Authors contributions
ACG co-developed the model, undertook the analyses and drafted the manuscript. CAD was involved in model development, contributed to the draft of the manuscript and provided estimates of BSE-infected animals. NMF co-developed the model and provided estimates of BSE-infected animals. RMA was involved in model development. All authors have read and approved the final manuscript.

Acknowledgements
This work was supported by the Department of Health. The views expressed in this publication are those of the authors and not necessarily those of the Department of Health. ACG and NMF acknowledge fellowship support from The Royal Society. NMF thanks the Howard Hughes Medical Institute for research funding, and RMA the Wellcome Trust. We thank Bob Will for providing data on the vCJD cases, and David Hilton and James Ironside for data on the appendix study.

References
prion protein genotype and sporadic Creutzfeldt-Jakob disease
Lancet 1999, 353:1673-1674
32. Department of Health Chief medical officer welcomes vCJD research results 2002.

Pre-publication history
The pre-publication history for this paper can be accessed here:

http://www.biomedcentral.com/1471-2334/3/4/prepub