Factors determining the pattern of the variant Creutzfeldt–Jakob disease (vCJD) epidemic in the UK

Azra C. Ghani*, Neil M. Ferguson, Christl A. Donnelly
and Roy M. Anderson

*Department of Infectious Disease Epidemiology, Faculty of Medicine, Imperial College, Norfolk Place,
London W2 1PG, UK*

Following the emergence of a new variant of Creutzfeldt–Jakob disease (vCJD) 6 years ago, and the gradual rise in clinical cases, there has been increased speculation regarding the overall magnitude of this epidemic in Great Britain. In this paper, we explore the epidemiological factors and uncertainties determining the scale of this epidemic in light of the most recent data on reported vCJD mortality. Our results demonstrate that, while the magnitude of the uncertainty has decreased dramatically since 1996, it is still not possible to predict with any degree of accuracy the final magnitude of this epidemic, with the 95% confidence interval for future cases being from 10 to 7000 deaths. However, short-term projections show that it is unlikely that a dramatic increase in case numbers will be observed in the next 2–5 years (95% confidence interval for 2 years: 10–80 cases, for 5 years: 10–200 cases). The results confirm significant age-dependent susceptibility/exposure to infection, with the likelihood profile demonstrating that those aged between 10 and 20 years are at highest risk of infection. We also demonstrate how projections based on onset data may be substantially biased, and explore the sensitivity of results to assumptions concerning the exposure to bovine spongiform encephalopathy (BSE) and the incubation-period distribution.

Keywords: variant Creutzfeldt–Jakob disease; bovine spongiform encephalopathy; epidemiology; mathematical model; projections

1. INTRODUCTION

Following the widespread exposure of the UK population to bovine spongiform encephalopathy (BSE)-infected meat in the late 1980s and early 1990s (Wilesmith *et al.* 1991; Anderson *et al.* 1996), and the emergence of a new variant of Creutzfeldt–Jakob disease (vCJD) in 1996 (Will *et al.* 1996), there has been growing speculation about the potential scale of infection and the future spread of this new disease. By the end of 2002, 129 cases of vCJD had been diagnosed in the UK, and eight of the patients were still alive. In addition, three cases have been confirmed in France (with a further two probable cases in patients who remain alive) (<http://www.invs.sante.fr>) and one in the Republic of Ireland, confirming fears that exposure to BSE-infected material may have been more widespread than originally thought. The relatively small number of vCJD cases diagnosed to date and the many biological and epidemiological uncertainties surrounding this new disease make prediction of the future course of the epidemic difficult. Because of the long incubation period of the disease (which may be of the order of 15–30 years or longer) and our still-limited understanding of disease pathogenesis and the factors determining susceptibility, it is not possible to estimate the prevalence of infection in the population from the vCJD case data alone. Thus, in the absence of a sensitive and specific test for asymptomatic infection,

predictions of the potential scale of this epidemic must rely on statistical modelling that incorporates the many uncertainties surrounding the biology and epidemiology of the disease. A similar problem was encountered early in the HIV/AIDS epidemic, with various statistical methods and mathematical models of transmission used to gain a better understanding not only of the potential scale of any epidemic, but also, more importantly, of the key unknowns that needed to be resolved to reduce the uncertainty in epidemic predictions (Anderson *et al.* 1989).

Much progress has been made recently in understanding the biology of transmissible spongiform encephalopathy (TSE) diseases in general and vCJD in particular. For example, more sensitive methods are now available to detect the abnormal form of the prion protein (PrP^{Sc}) in different tissues (in particular, lymphoreticular tissues), potentially allowing the detection of infection prior to the onset of clinical symptoms (Hill *et al.* 1999; Maissen *et al.* 2001; Wadsworth *et al.* 2001). Three large-scale studies testing tonsil and appendix tissues are currently underway in the UK to determine the prevalence of asymptomatic infection (Hilton *et al.* 2002). In addition, future studies are planned in Switzerland (which experienced an epidemic of BSE in cattle *ca.* 3 years later than the UK; Doherr *et al.* 1999) and the Republic of Ireland. These studies will provide much-needed estimates of the infection prevalence in the various populations, although the unknown sensitivity and specificity of the tests at different stages of the incubation period will necessitate careful interpretation of any results. In addition, it is not currently

* Author for correspondence (a.ghani@imperial.ac.uk).

known what proportion of infected individuals will develop clinical disease. In particular, such studies will provide the only means by which the potential for ongoing secondary transmission of infection from human to human (for example, through surgical instruments or via blood products) can be assessed.

We explore how the predictability of the vCJD epidemic has changed over the past 4 years and the epidemiological factors that will determine the future course of the epidemic.

2. vCJD EPIDEMIOLOGICAL DATA

By the end of December 2001, 113 cases of vCJD had been diagnosed in the UK, and nine of the patients remained alive (<http://www.doh.gov.uk/cjd>). A further three cases had been confirmed in France and one in the Republic of Ireland. All cases in the UK are defined based on a set of diagnostic criteria (including progressive neuropsychiatric disorder) and are confirmed only at post-mortem by neuropathology (Will *et al.* 2000). From mid-1999 onwards, the epidemiological statistics were widened to include 'probable' cases defined on the basis of a set of diagnostic criteria and, where available, a positive tonsil biopsy, these 'probable' cases are included as such until definite diagnosis is made at post-mortem. A small number of cases are never confirmed pathologically because a post-mortem examination does not take place, but are included in the case statistics. However, all probable cases that have had a post-mortem examination have been confirmed neuropathologically.

Figure 1*a* shows the time-series of the UK cases stratified by year of onset of symptoms (including all probable cases) and year of death (including probable cases awaiting post-mortem confirmation). There is, in general, a delay between the onset of clinical symptoms and the reporting of a probable case, with a median duration of 13 months and a maximum of 39 months. The reported onsets for 2000 and 2001 (as reported by the end of 2001) are therefore almost certainly incomplete. Figure 1*a* also shows the reported number of onsets for 2000 adjusted based on the reporting delay for 1999. As only one case had been reported by the end of 2001 with onset in 2001, we consider the onsets only to the end of 2000. By contrast, there is no significant delay in reporting for deaths since the majority of cases now enter the statistics as probable cases while the patient is still alive.

Figure 1*b* shows the age distribution at onset and death of the UK cases. The median age is 26 years at the onset of symptoms and 28 years at death. The substantial number of younger cases indicates that younger individuals are more susceptible to infection, or were exposed to a greater extent and/or have shorter incubation periods (Ghani *et al.* 2000*a*). The age distribution of the cases has remained remarkably stable over the past 5 years, with little change in the median age either at onset of symptoms or at death.

To date, all of the tested cases ($n=87$) have shared a common genetic trait, being methionine homozygous (MM) at codon 129 of the prion protein (PrP) gene (The National CJD Surveillance Unit 2000). Studies in Caucasian populations estimate that *ca.* 40% of the population share this trait, with 13% being valine homozygous (VV) and the remaining 47% heterozygous (MV) (Owen *et al.*

1990; Collinge *et al.* 1991). A more recent study has demonstrated a decreased risk of vCJD in those with HLA-DQ7 type (Jackson *et al.* 2001). While it is not at present known whether there is any correlation between codon 129 genotype and this HLA type, these new results suggest complex genetic (i.e. multigene) determinants of patterns of susceptibility and potentially pathogenesis for this class of diseases.

Spatial analyses indicate that the incidence of vCJD in the UK is significantly higher in the north than in the south (Cousens *et al.* 2001). These analyses have also identified a geographical cluster of five cases of vCJD, all of which could be linked to the village of Queniborough in Leicestershire (Cousens *et al.* 2001). A detailed local investigation into this cluster revealed that four of the patients had consumed beef purchased from one of two local butchers during the early 1980s (Bryant & Monk 2001). Both of these butchers slaughtered animals on site using traditional practices and hence it is possible that meat sold by these butchers could have been infected with the BSE agent through contact with brain or other central-nervous-system tissue. A test of significance indicated that these patients were significantly more likely to have consumed meat purchased from one of these higher-risk butchers prior to 1986 than were controls (Bryant & Monk 2001), potentially therefore providing some information on the incubation period of the disease in this small number of cases.

3. MODEL FRAMEWORK AND STATISTICAL METHODS

A time- and age-stratified mathematical model of the transmission dynamics of vCJD is used to relate the exposure of the population to BSE-infected meat products to the observed incidence of vCJD. Given the current lack of data with which to estimate the prevalence of asymptomatic infection, we restrict our analyses to consideration of primary transmission from cattle to humans, and do not consider the theoretical risks of secondary transmission from humans to humans (for example, through blood products or via surgical instruments). Full details of this model have previously been published (Ghani *et al.* 1998*a*, 2000*b*). In brief, the probability that an individual develops clinical disease (defined by either onset of symptoms or death) at time u and age a is given by

$$p(u, a) = S(a) \int_{u-a}^u f(u-t) \beta I(t, a-u+t) \times \exp \left[- \int_0^t \beta I(t', a-u+t') dt' \right] dt, \quad (3.1)$$

where $S(a)$ is the survival probability, $f(u)$ is the incubation-period distribution and β is the transmission coefficient. The exposure hazard $I(t, a)$ is given by

$$I(t, a) = \nu(t) g(a) \int \Omega(z) w(z, t) dz. \quad (3.2)$$

Here, $\nu(t)$ is the effectiveness of control measures (namely the specified bovine offal ban in 1989, which is allowed to vary between being 0 and 100% effective), $g(a)$ is an age-dependent susceptibility/exposure function, $\Omega(z)$ is the relative infectiousness of bovines over the course of their incubation period and $w(z, t)$ is the proportion of cattle slaughtered at time t and

Figure 1. (a) Time-series of 113 vCJD onsets (grey bars) and 104 deaths (filled bars) as reported by 1 January 2002. Onsets are adjusted for reporting delay (open bars); onsets in 2000 are adjusted using estimates obtained from onsets in 1999. (b) Age distribution of vCJD onsets (light grey bars) and deaths (dark grey bars) as reported by 1 January 2002. (c) Pattern of exposure to BSE-infected material stratified by incubation stage (number of cattle slaughtered per quarter year of calendar time, and incubation period). The estimates of numbers of infected animals within six months of onset of clinical signs include clinical cases diagnosed before mid-1988 and estimates of non-reported clinical cases, obtained from updated analyses of BSE case data to August 2001. (d) The mean number of months between onset of clinical symptoms of vCJD and death, by year of onset.

time z away from disease onset. This model is equivalent to a more complex dose-dependent model (Ghani *et al.* 1998a) if we assume a linear dose-response curve (Diringer *et al.* 1998; McLean & Bostock 2000). This simplification considerably reduces the computational complexity of the model.

The expected number of cases (onsets or deaths) at time u and age a is given by

$$c(u, a) = r(u)B(u - a)p(u, a), \quad (3.3)$$

where $r(u)$ is the probability that a case arising at time u is reported and $B(u - a)$ is the number of individuals born at time $u - a$. When fitting to the vCJD deaths, we assume that $r(u) = 1$ for all values of u . For vCJD onsets, we assume that the reporting delay for cases with onsets in 2000 is similar to that for cases with onsets in 1999, and therefore set $r(2000) = 0.84$ and $r(u) = 1$ otherwise.

The model is fitted using maximum-likelihood methods to the time- and age-stratified vCJD onsets or deaths, assuming that the data arise from a Poisson distribution. The individual cases are categorized into cells with at least five observations in each. The best-fit point is that which maximizes the log likelihood (LnL), or equivalently minimizes

$$-2\text{LnL} = 2 \sum_{u,a} c(u, a) - x(u, a) \ln(c(u, a)) - \ln(x(u, a)!), \quad (3.4)$$

where $x(u, a)$ is the observed number of cases at time u and age a . For any model parameter, maximum-likelihood confidence intervals are obtained using likelihood ratio tests. Nonlinear optimization techniques are used to fit the model (Press *et al.* 1992) 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.

To enable confidence intervals to be calculated for the projected number of cases, it was necessary to re-parameterize the model such that the expected number of cases between times u_1 and u_2 , $T(u_1, u_2)$, became a model parameter. Combining equations (3.1) and (3.3), $T(u_1, u_2)$ is given by

$$T(u_1, u_2) = \int_{u_1}^{u_2} \int_a r(u)B(u - a)S(a) \int_{u-a}^u f(u - t, a - u + t) \times \beta I(t, a - u + t) \exp \left[- \int_0^t \beta I(t', a - u + t') dt' \right] dt da du. \quad (3.5)$$

If infection incidence is very low within any given cohort, the saturation factor for the hazard function can be ignored, so that equation (3.5) simplifies to give:

$$c(u_1, u_2) = \int_{u_1}^{u_2} \int_a r(u) N(a) S(u, a) \int_{u-a}^u f(u-t, a-u+t) \times \beta I(t, a-u+t) dt da du. \quad (3.6)$$

However, in general, given the uncertainty in the scale of the vCJD epidemic (and that many scenarios that are consistent with the vCJD case data have saturation within cohorts), approximation (3.6) is not valid. For any proposed parameter set, equation (3.5) was therefore solved numerically to calculate β . If no solution was possible (for example, if the proposed parameter set generated fewer infections than the required epidemic size) then the proposed point in the optimization routine was rejected.

The results from the model are sensitive to the assumptions made regarding the exposure to BSE-infected material, $w(z, t)$. In our early work, we assumed that this exposure is proportional to estimates of the numbers of BSE-infected preclinical cattle that entered the British food supply from 1980 onwards (Ghani *et al.* 1998*a*, 2000*b*). However, prior to BSE becoming a notifiable disease in 1988, it is possible that clinical cases (both those observed and those not reported) may have entered the food supply. Sensitivity analyses have demonstrated that inclusion of estimates of the number of BSE cases prior to 1988 does not substantially alter short- and long-term projections of the future magnitude of the epidemic (Ferguson *et al.* 2002). Our analyses are therefore based on estimates of exposure (using British BSE case data to August 2001) to preclinical infected animals, numbers of clinical BSE cases reported prior to mid-1988 and estimates of unreported cases (see figure 1*c*).

Other recent studies estimating the potential scale of the vCJD epidemic in Great Britain have based exposure on BSE clinical cases (using either the actual number of cases (Huillard d'Aignaux *et al.* (2001)), or a simplified exponential function (Valleron *et al.* (2001))). The use of reported BSE cases is a poor proxy measure for exposure to BSE since there is likely to have been substantial early underreporting of BSE cases. In a later section, we explore the sensitivity of results to the assumptions made regarding exposure to infectious cattle.

Pathogenesis experiments in cattle have indicated that BSE infectivity is most widely distributed in the carcass towards the end of the incubation period (Wells *et al.* 1994, 1998). 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. Clinical BSE cases (including those reported prior to mid-1988 and unreported cases) are assumed to be as infectious as those preclinical animals within six months of the onset of disease. While the parameter for this exponential increase is fitted in the model, the results are fairly insensitive to its value.

The stable age distribution of the cases suggests that younger individuals are either more susceptible to infection or were exposed to a greater extent and that it is unlikely that this age distribution arose from an age-dependent incubation period alone (Ghani *et al.* 2000*a*). This is because, under the latter hypothesis, we would expect an increasing age distribution over time, and hence models without age-dependent susceptibility or exposure do not provide good fits to the data. We explored a number of different parametric forms for the age-dependent susceptibility/exposure function, including uniform, normal and

gamma distributions. The best-fitting distributional form is a three-parameter combined uniform distribution with gamma-distributed tails, which, at its limits, is either the uniform or the 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 generally long (in the order of 10–40 years plus) and highly variable (Huillard d'Aignaux *et al.* 1999, 2002; Brown *et al.* 2000). For many TSEs, including BSE, there is a minimum incubation period so that functional forms must allow for a substantial delay following exposure to infection prior to the onset of clinical symptoms of vCJD (Asher *et al.* 1973; Anderson *et al.* 1996). Since the standard statistical survival distributions (including the Weibull, gamma and lognormal distributions) have restricted higher moments (skewness and kurtosis), we use a modified form of the four-parameter generalized lambda distribution (Ramberg *et al.* 1979) (with inverse cumulative distribution function $X(p) = \lambda_1 + (p^{\lambda_3} + (1-p)^{-\lambda_4} - 1)/\lambda_2$). This distribution is more flexible than distributions based on standard functional forms, and incorporates all of the standard functional survival distributions as special cases. We also explore the sensitivity of results to assumptions regarding the functional form of the incubation-period distribution.

Since vCJD cases have been diagnosed only in those who are MM-homozygous at codon 129 of the PrP gene, we restrict our analyses to this proportion of the population. Those with other genotypes may also develop disease in the future; however, we are unable to constrain scenarios in these other genotypes and hence the epidemic remains unpredictable in these genotypes. However, if other genetic characteristics also influence susceptibility to infection, it is possible that the size of the susceptible population may be smaller than the 40% assumed. We therefore examine the sensitivity of our results to this proportion.

4. PREDICTABILITY OF THE MAGNITUDE OF THE EPIDEMIC

The long-term scale of the vCJD epidemic remains uncertain, with wide confidence bounds obtained for projections to 2020, 2040 and 2080 (table 1). The 95% upper confidence bound on these projections varies substantially depending on whether the model is fitted to the onset data or to deaths, with an upper value of 40 000 (for vCJD cases to 2080) obtained in the former case and 7000 in the latter. In the absence of any treatments that improve survival, as is currently the case, the time-series for deaths is not influenced by changes in awareness or reporting of the disease, although it may be susceptible to other biases, such as families seeking medical support to keep their relatives alive. By contrast, the onset data can potentially be influenced by increased awareness of the symptoms of the disease, which is likely to result in earlier reporting of the onset of symptoms. Figure 1*d* shows the variation in the time between onset of symptoms and death by year of onset of symptoms. This variation, which may be a consequence of variability in the infection process and/or indicate changes in reporting, results in a difference in the time-series for onsets and deaths, which in turn may bias projections based on onset data. A further complication in fitting to onset data is the delay in reporting. For example, 20% of onsets in 1999 were reported in the same year, 64% the following year, and the remaining 16% in

Table 1. Projections of the magnitude of the vCJD epidemic in Great Britain.

(Projections are based on fitting to either the vCJD onsets to the end of 2000 (correcting for reporting delay) or deaths to the end of 2001. The MLE denotes the maximum-likelihood estimate; CI denotes the confidence interval. Maximum-likelihood parameter estimates are available from the authors on request.)

long-term projections		to 2020	to 2040	to 2080
deaths (from 2002)	MLE	70	80	80
	95% CI	10–1500	10–4000	10–7000
onsets (from 2001)	MLE	800	2000	3000
	95% CI	10–4000	10–20 000	10–40 000
deaths assuming three cases infected prior to 1986	MLE	20	20	20
	95% CI	10–1500	10–5000	10–8000
sensitivity to exposure function deaths based on BSE cases	MLE	90	90	100
	95% CI	10–1500	10–4000	10–6000
sensitivity to incubation-period distribution deaths based on lognormal distribution	MLE	125	125	125
	95% CI	30–500	30–500	30–500
deaths based on offset gamma distribution	MLE	70	70	70
	95% CI	30–1500	30–3000	30–6000
deaths based on offset lognormal distribution	MLE	80	80	80
	95% CI	30–1000	30–2000	30–3000
deaths based on offset <i>F</i> distribution	MLE	90	90	90
	95% CI	10–700	10–1500	10–1500
short-term projections		2 year	5 year	10 year
deaths (from 2002)	MLE	30	70	90
	95% CI	10–80	10–200	10–600
onsets (from 2001)	MLE	60	175	400
	95% CI	10–100	10–400	10–1000

2001. To make projections based on onset data, it is therefore important to include some adjustment for this reporting delay (Andrews *et al.* 2000). It is likely, however, that this reporting delay may change over time and hence projections based on onset data may be biased. Therefore, while the onset data remain important since they provide earlier information on the pattern of the epidemic, current projections based on onset data may be severely biased.

Figure 2*a–c* shows the likelihood profile for the long-term projections based on fitting to deaths. The maximum-likelihood estimate (MLE) for future case numbers is small (80 cases to 2080). However, the relatively flat profile demonstrates that maximum-likelihood estimates for vCJD projections must be interpreted with caution, since many scenarios with a much higher number of cases also fit the data and are well within the 95% confidence interval (10–7000 deaths). In this context, it is also important to note that any future data that exclude epidemic scenarios in the vicinity of the current MLE may result in increases in the upper 95% confidence bound on epidemic size.

While these results continue to demonstrate great uncertainty in the future magnitude of the epidemic, they also demonstrate a substantial reduction in this uncertainty over the past 4 years. Figure 2*d* shows the 95% confidence intervals on cases to 2080 obtained by fitting to the case data (deaths) to the end of 1998, 1999, 2000 and 2001. In our previous analyses (Ghani *et al.* 1998*a*, 2000*a,b*), ranges of epidemic sizes consistent with the data

were based on scenario analyses and hence are not directly comparable to the results presented here and in more recent analyses (Ferguson *et al.* 2002). Although the trend of decreasing uncertainty is the same, and a similar order of magnitude in ranges is obtained, the scenario analyses used in earlier publications (Ghani *et al.* 1998*a*, 2000*a,b*) give slightly wider bounds. It is not possible, *a priori*, to predict which of the two different statistical methods would give wider bounds, and different results may be obtained for other problems. When the data are very limited (as was the case early in the epidemic), fitting methods are inappropriate as similar fits can be obtained from a large number of different scenarios, and hence identification of the best-fitting model is difficult. However, when sufficient data become available, maximum-likelihood fitting methods are more efficient and allow estimation of parameters of interest (such as future case numbers and the mean incubation period).

Over the past 4 years, the upper 95% confidence bound on future cases to 2080 has decreased from approximately half a million cases (representing very high uncertainty) to 7000 cases. This decrease largely results from the observed epidemic curve being less steep than would be required to generate the largest epidemic scenarios. However, the confidence interval has also been informed by the stable age distribution of cases, which suggests that there are restricted cohorts of younger individuals who are more susceptible to infection and/or were exposed to a greater extent (Ghani *et al.* 2000*a*).

Figure 2. Univariate-likelihood profiles for the number of vCJD deaths from (a) 2002–2080, (b) 2002–2040 and (c) 2002–2020. The bold lines show the cut-off corresponding to the 95% confidence limit (based on the maximized log-likelihood and one degree of freedom). (d) 95% confidence intervals for the number of vCJD deaths to 2080 based on maximum-likelihood fitting of the model to the vCJD deaths confirmed to the end of 1998, 1999, 2000 and 2001.

Investigation of the geographical cluster of cases in Queniborough, Leicestershire, suggested additional information on the incubation period of three cases of vCJD (Bryant & Monk 2001). By constraining the model to be consistent with these three cases becoming infected prior to 1986, the upper 95% confidence interval on the long-term scale of the epidemic is marginally increased to 8000 cases.

Table 1 shows the projections of numbers of cases in the next 2, 5 and 10 years. The results demonstrate the slow course of any epidemic, such that even in the worst-case scenarios, it is now unlikely that a very large increase in case numbers will be seen in the short-term.

5. PATTERN OF EXPOSURE TO INFECTIOUS CATTLE

Projections of the future course and pattern of the vCJD epidemic rely on the pattern of exposure to BSE-infected cattle in the 1980s and early 1990s, without which it would be impossible to make any projections. Figure 1c shows the pattern of exposure to BSE cases and infected animals, with the number of the latter based on updated estimates obtained from a back-calculation model for BSE

(Anderson *et al.* 1996; Ferguson *et al.* 1997). In previous analyses, we assumed that exposure was proportional to the number of preclinical infected animals entering the food supply. However, it is also possible that BSE clinical cases entered the food supply prior to BSE becoming a notifiable disease in 1988. Inclusion (or exclusion) of these animals in the exposure function does not alter either the likelihood profiles obtained or the 95% confidence intervals on future epidemic size.

Three recent papers (Cousens *et al.* 1997; Huillard d'Aignaux *et al.* 2001; Valleron *et al.* 2001) projecting the future scale of the vCJD epidemic have used cruder assumptions about exposure, based on the pattern of BSE clinical cases reported and not accounting for the substantial number (*ca.* 750 000 animals prior to 1996; Anderson *et al.* 1996; Ferguson *et al.* 1997) of pre-clinical infected animals that entered the food supply. While these cruder assumptions influenced the results obtained in the past, with narrower confidence intervals obtained on long-term projections, the results are now less sensitive to the assumptions made regarding exposure, suggesting that more information is now being derived from the pattern of vCJD deaths.

6. EPIDEMIOLOGICAL DETERMINANTS OF THE COURSE OF THE EPIDEMIC

(a) Incubation-period distribution

In the absence of knowledge of the protective effect of the species barrier and thus the infectious dose for humans, the incubation period of the disease becomes the major determinant of the future course of the epidemic (Cousens *et al.* 1997; Ghani *et al.* 1998a). Some parallels can be drawn with other human TSEs, such as sporadic CJD and kuru, which demonstrate long and variable incubation periods. However, direct comparison cannot be made since oral transmission (compared with intracerebral for iatrogenic CJD) and the species barrier (which is absent for kuru) have both been demonstrated to lengthen and increase variability in the incubation period in animal experiments (Kimberlin & Walker 1978; Diringer *et al.* 1998).

Figure 3a shows the univariate-likelihood profile for the mean incubation period. This demonstrates that the case data to the end of 2001 do not allow any estimate of the upper bound on this parameter to be made. However, the delay between exposure (which peaked in the late 1980s–early 1990s) and the appearance of the first clinical cases in 1994–1995 results in a lower 95% confidence bound of 7 years. Including the additional restriction of at least three cases infected prior to 1986 (as suggested by the Queniborough cluster) increases this minimum bound to 10 years.

The relationship between the mean incubation period and the magnitude of the epidemic is illustrated in figure 3b, which shows the bivariate-likelihood profile for mean incubation period and vCJD deaths to 2080. If the incubation period is relatively short (in the region of 10–20 years), then the confidence interval for future cases is narrow. By contrast, for longer mean incubation periods, the potential scale of the epidemic is both larger and more variable. It is important to note that we are assuming that only individuals who are MM at codon 129 of the PrP gene are at risk of developing vCJD. Other genotypes may have longer incubation periods, as suggested by recent studies of kuru (Huillard d'Aignaux *et al.* 1999, 2002; Brown *et al.* 2000).

Projections of the long-term scale of the epidemic are also highly sensitive to the parametric form assumed for the incubation-period distribution. Clearly, if the incubation-period distribution is multimodal (as could potentially occur with wide-scale genetic variation in susceptibility), then the scale of the epidemic could be much greater than presented here. However, restricting analysis to the standard survival distributions with more restricted higher moments (including the gamma, lognormal and *F* distributions; Cousens *et al.* 1997; Huillard d'Aignaux *et al.* 2001; Valleron *et al.* 2001) in general results in narrower 95% confidence bounds on projections (see table 1), in addition to influencing the confidence intervals obtained for other parameters (such as the effectiveness of the specified bovine offal ban).

(b) Age-dependent susceptibility/exposure to infection

As demonstrated in previous analyses, the age distribution of the vCJD cases provides important information on the pattern of the epidemic. The young age distribution

Figure 3. (a) Univariate-likelihood profile for the mean incubation period. The bold line shows the cut-off corresponding to the 95% confidence limit (based on the maximized log-likelihood and one degree of freedom). (b) Bivariate-likelihood profile for mean incubation period and deaths to 2080. The red region is the best-fitting model. Colours denote the *p*-value for the two-dimensional confidence region, with the 0.05 boundary representing the 95% confidence region.

of cases could not have arisen without some form of age-dependent susceptibility and/or exposure to infection. In most epidemiological studies of infectious diseases it is not possible to disentangle the relative contributions of susceptibility and exposure on the basis of reported age distributions of disease incidence, unless additional independent information (such as dietary patterns, which are currently being studied) is available on either variable. Figure 4 shows the maximum-likelihood estimate of the pattern of age-dependent susceptibility/exposure to infection. The results demonstrate that those between *ca.* 10 and 20 years of age are most susceptible/exposed to infection, with those over 40 years old being much less susceptible/exposed. This pattern of age-dependent susceptibility/exposure restricts the scale of the epidemic by restricting the extent of infection in older individuals. However, even for smaller epidemics, it does not totally

Figure 4. Maximum-likelihood estimate of the age-dependent relative susceptibility/exposure function.

Figure 5. Univariate-likelihood profile for the effectiveness of the specified bovine offal ban introduced in June 1989. The bold line shows the cut-off corresponding to the 95% confidence limit (based on the maximized log-likelihood and one degree of freedom).

exclude the possibility of cases arising in older individuals, as evidenced by the case diagnosed in a 74-year-old man (Lorains *et al.* 2001).

The pattern of age-dependent susceptibility/exposure is important to the design and interpretation of large-scale unlinked anonymous testing programmes for the prevalence of infection, particularly if the tests are less sensitive early in the course of infection. While a random survey of the population is necessary to provide an estimate of the prevalence of infection in different cohorts, surveys targeting those at greatest risk are most likely to obtain a positive result and hence reduce the current uncertainty in the scale of the epidemic (Ghani *et al.* 1998*b*, 2000*a*).

(c) Effectiveness of the specified bovine offal ban

The specified bovine offal ban introduced in mid-1989 was designed to remove the potentially most infectious material (including brain, spinal cord and other central nervous tissue) from the human food supply. However, because of difficulties in establishing compliance with this order and because additional tissues, such as dorsal root ganglia, later shown to infect mice were not banned in 1989, it is likely that this measure was not 100% effective in removing infectious material from the food chain. Figure 5 shows the likelihood profile obtained for this parameter. While the maximum-likelihood estimate for this parameter (*ca.* 95%) suggests that this ban was highly

Figure 6. Univariate-likelihood profiles for the number of vCJD deaths to 2080 for varying assumptions regarding the proportion of the population that is genetically susceptible to infection. The bold line shows the cut-off corresponding to the 95% confidence limit (based on the maximized log-likelihood and one degree of freedom). 5% susceptible, filled triangles; 10% susceptible, grey squares; 40% susceptible, grey diamonds.

effective, the 95% confidence interval ranges from 0% to 100% effective, demonstrating that epidemiological analysis is unable, at the current time, to provide any information on the effectiveness of the specified bovine offal ban.

(d) Genetic susceptibility to infection

Recent studies have indicated that there may be substantial genetic variation in susceptibility to TSE infection and disease pathogenesis (Windl *et al.* 1996; Jackson *et al.* 2001; Lloyd *et al.* 2001). In the absence of detailed information on the relative susceptibility to infection of those with different genetic characteristics together with population-based surveys of these characteristics, it is not at present possible to explore, in detail, the influence of this variation on projections. However, figure 6 shows the likelihood profiles obtained assuming that the cases arose from different proportions of the population (5, 10 and 40%). The results show that the projections of the scale of the epidemic are robust to this level of variation in susceptibility.

7. CONCLUSIONS

The magnitude of the vCJD epidemic in the UK remains uncertain, with wide confidence intervals obtained for long-term projections of future case onsets and deaths. However, the relatively low annual incidence of cases recorded over the past 5 years has substantially reduced the previous uncertainty in the scale of the epidemic in the known genetically susceptible population, with a reduction in the upper 95% confidence interval from approximately half a million deaths, based on projections made in 1998, to 7000 based on deaths to the end of 2001. Shorter-term projections of the scale of the epidemic currently exclude the possibility of large numbers of cases in the next 2–5 years. However, the many uncertainties in the epidemiology and biology of this disease, which justify only a minimal set of assumptions to underpin the projections, mean that these numbers should be interpreted as representing the range of uncertainty rather than as absolute bounds on epidemic size. In particular, our results demonstrate that a wide range of epidemic

scenarios are consistent with the vCJD case data to the end of 2001, such that 'best' estimates of future case numbers must be interpreted with caution.

As for other long-incubation-period diseases (Isham 1989; Ferguson *et al.* 1997), in the absence of independent data on the incubation-period distribution or the infection rate, it is difficult either to predict accurately the scale of the epidemic or to estimate the mean incubation period from the observed case data alone while the epidemic is in its early stages. However, for vCJD it is possible to infer epidemic patterns consistent with the case data based on the pattern of exposure to BSE-infected material in the 1980s and early 1990s. Recent projections for vCJD in the UK by other groups have been based on much simplified assumptions regarding both the exposure pattern and the form of the incubation period distribution. Although the current results are less sensitive to the exposure pattern than in past analyses, they remain sensitive to the form of the incubation-period distribution. In particular, not allowing for the initial delay (Valleron *et al.* 2001) observed for many TSEs cannot be justified on the basis of currently available data and substantially reduces the range of scenarios that appear consistent with the case data.

As in previous work, our results demonstrate the importance of the length of the incubation period in determining the future course of any epidemic, with longer incubation periods resulting in larger epidemics. In contrast to other studies (Valleron *et al.* 2001) based on fitting to the same data, we were unable to estimate the mean incubation period. Indeed, unless the number of vCJD cases remains at a low level, it will be difficult precisely to estimate this quantity without independent data on infection rates (obtained, for example, through prevalence surveys). However, by relating the pattern of exposure to BSE-infected material to the time-series of vCJD cases, it is possible to obtain a meaningful lower 95% confidence interval for the mean incubation period, which we estimate to be *ca.* 7 years. This lower 95% confidence interval is increased to *ca.* 10 years by additionally conditioning on three of the current vCJD cases having been infected prior to 1986, as suggested by detailed research into the cluster of vCJD cases in Queniborough, Leicestershire (Bryant & Monk 2001).

Our analyses confirm previous findings regarding age-dependent susceptibility/exposure to infection (Ghani *et al.* 2000a), with those aged 10–20 years at highest risk, and those aged over *ca.* 40 years at much lower risk. However, it should be noted that a wide range of age-dependent susceptibility/exposure patterns remain consistent with the data. Again, given the limited number of vCJD cases that have arisen to date, our analyses suggest that it is not at present possible to infer any more precisely the degree of age-dependent susceptibility/exposure, as suggested by other studies (Valleron *et al.* 2001). However, the patterns observed remain important to the design of large-scale unlinked anonymous testing programmes to determine the prevalence of asymptomatic infection (Ghani *et al.* 2000a), particularly if such programmes use routinely removed tissues (such as tonsil tissues) that are not a random sample of the population of interest.

Unless the numbers of vCJD cases in the next few years show a significant decline, it is likely that predictions of

the primary epidemic based on case data alone will remain uncertain. In addition, such analyses are unable to provide any indication of the potential for secondary transmission of the agent, for example through blood products or via surgical instruments. Large-scale studies are currently underway to test tonsil and appendix tissues for the presence of PrPSc (Hilton *et al.* 2002). Although the results of these studies will provide useful information on the prevalence of late-stage infection, their interpretation will be limited by our relatively poor understanding of the sensitivity and specificity of the 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 main priority therefore remains the development of a diagnostic test that is able to detect infection early in the incubation period and that can be applied to large population samples, both in the UK and elsewhere.

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. A.C.G. and N.M.F. acknowledge fellowship support from The Royal Society. N.M.F. thanks the Howard Hughes Medical Institute for research funding, and R.M.A. thanks the Wellcome Trust. The authors thank Bob Will at the CJD Surveillance Unit for providing data on the vCJD cases.

REFERENCES

- Anderson, R. M., Cox, D. R. & Hillier, H. C. 1989 Epidemiological and statistical aspects of the AIDS epidemic: introduction. *Phil. Trans. R. Soc. Lond. B* **325**, 1–6.
- Anderson, R. M. (and 14 others) 1996 Transmission dynamics and epidemiology of BSE in British cattle. *Nature* **382**, 779–788.
- Andrews, N. J., Farrington, C. P., Cousens, S. N., Smith, P. G., Ward, H., Knight, R. S. G., Ironside, J. W. & Will, R. G. 2000 Incidence of variant Creutzfeldt–Jakob disease in the UK. *Lancet* **356**, 481–482.
- Asher, D. M., Gibbs, C. J., David, E., Alpers, M. P., Gajdusek, D. C. & Sadowsky, D. A. 1973 Experimental kuru in the chimpanzee: physical findings and clinical laboratory studies. In *Symp. IVth Int. Congr. Primatology Nonhuman primates and human diseases* (ed. W. P. McNulty), pp. 43–90. Basel: Karger.
- Brown, P. (and 14 others) 2000 Iatrogenic Creutzfeldt–Jakob disease at the millennium. *Neurology* **55**, 1075–1081.
- Bryant, G. & Monk, P. 2001 Summary of the final report of the investigation into the North Leicestershire cluster of variant Creutzfeldt–Jakob disease. Leicestershire NHS Health Authority.
- Collinge, J., Palmer, M. S. & Dryden, A. J. 1991 Genetic predisposition to iatrogenic Creutzfeldt–Jakob disease. *Lancet* **337**, 1441–1442.
- Cousens, S. N., Vynnycky, E., Zeidler, M., Will, R. G. & Smith, P. G. 1997 Predicting the CJD epidemic in humans. *Nature* **385**, 197–198.
- Cousens, S. (and 10 others) 2001 Geographical distribution of variant Creutzfeldt–Jakob disease in Great Britain, 1994–2000. *Lancet* **357**, 1002–1007.
- Diringer, H., Roehmel, J. & Beekes, M. 1998 Effect of repeated oral infection of hamsters with scrapie. *J. Gen. Virol.* **79**, 609–612.
- Doherr, M. G., Heim, D., Vandeveld, M. & Fatzer, R. 1999 Modelling the expected number of preclinical and clinical cases of bovine spongiform encephalopathy in Switzerland. *Vet. Rec.* **145**, 155–160.

- Ferguson, N. M., Donnelly, C. A., Woolhouse, M. E. J. & Anderson, R. M. 1997 The epidemiology of BSE in cattle herds in Great Britain. II. Model construction and analysis of transmission dynamics. *Phil. Trans. R. Soc. Lond. B* **352**, 803–838. (DOI 10.1098/rstb.1997.0063.)
- Ferguson, N. M., Ghani, A. C., Donnelly, C. A., Hagenaars, T. J. & Anderson, R. M. 2002 Estimating the risk to human health posed by possible entry of BSE infection into the GB sheep flock. *Nature* **415**, 420–425.
- Ghani, A. C., Ferguson, N. M., Donnelly, C. A., Hagenaars, T. J. & Anderson, R. M. 1998a Epidemiological determinants of the pattern and magnitude of the vCJD epidemic in Great Britain. *Proc. R. Soc. Lond. B* **265**, 2443–2452. (DOI 10.1098/rspb.1998.0596.)
- Ghani, A. C., Ferguson, N. M., Donnelly, C. A., Hagenaars, T. J. & Anderson, R. M. 1998b Estimation of the number of people incubating variant CJD. *Lancet* **352**, 1353–1354.
- Ghani, A. C., Donnelly, C. A., Ferguson, N. M. & Anderson, R. M. 2000a Assessment of the prevalence of vCJD through testing tonsils and appendixes for the abnormal prion protein. *Proc. R. Soc. Lond. B* **267**, 23–29. (DOI 10.1098/rspb.2000.0961.)
- Ghani, A. C., Ferguson, N. M., Donnelly, C. A. & Anderson, R. M. 2000b Predicted vCJD mortality in Great Britain. *Nature* **406**, 583–584.
- Hill, A. F. (and 14 others) 1999 Investigation of variant Creutzfeldt–Jakob disease and other human prion diseases with tonsil biopsy samples. *Lancet* **353**, 183–189.
- Hilton, D. A., Ghani, A. C., Conyers, L., Edwards, P., McCardle, L., Penney, M., Ritchie, D. & Ironside, J. W. 2002 Accumulation of prion protein in tonsil and appendix: review of tissue samples. *Br. Med. J.* **325**, 633–634.
- Huillard d'Aignaux, J. (and 10 others) 1999 Incubation period of Creutzfeldt–Jakob disease in human growth hormone recipients in France. *Neurology* **53**, 1197–1201.
- Huillard d'Aignaux, J. N., Cousens, S. N. & Smith, P. G. 2001 Predictability of the UK variant Creutzfeldt–Jakob disease epidemic. *Science* **294**, 1729–1731.
- Huillard d'Aignaux, J. N., Cousens, S. N., Maccario, J., Costagliola, D., Alpers, M. P., Smith, P. G. & Alperovitch, A. 2002 An analysis of the incubation period of kuru by sex and age at infection. *Epidemiology* **15**, 402–408.
- Isham, V. 1989 Estimation of the incidence of HIV infection. *Phil. Trans. R. Soc. Lond. B* **325**, 113–121.
- Jackson, G. S., Beck, J. A., Navarrete, C., Brown, J., Sutton, P. M., Conteras, M. & Collinge, J. 2001 HLA-DQ7 antigen and resistance to variant CJD. *Nature* **414**, 269–270.
- Kimberlin, R. H. & Walker, C. A. 1978 Pathogenesis of mouse scrapie: effect of route of inoculation on infectivity titres and dose–response curves. *J. Comp. Pathol.* **88**, 39–47.
- Lloyd, S. E., Onwuazor, O. N., Beck, J. A., Mallinson, G., Farrall, M., Targonski, P., Collinge, J. & Fisher, E. M. C. 2001 Identification of multiple quantitative trait loci linked to prion disease incubation period in mice. *Proc. Natl Acad. Sci. USA* **98**, 6279–6283.
- Lorains, J. W., Henry, C., Agbamu, D. A., Rossi, M., Bishop, M., Will, R. G. & Ironside, J. W. 2001 Variant Creutzfeldt–Jakob disease in an elderly patient. *Lancet* **357**, 1339–1340.
- McLean, A. R. & Bostock, C. J. 2000 Scrapie infections initiated at varying doses: an analysis of 117 titration experiments. *Phil. Trans. R. Soc. Lond. B* **355**, 1043–1050. (DOI 10.1098/rstb.2000.0641.)
- Maissen, M., Roeckl, C., Glatzel, M., Goldmann, W. & Aguzzi, A. 2001 Plasminogen binds to disease-associated prion protein of multiple species. *Lancet* **357**, 2026–2028.
- Owen, F., Poulter, M., Collinge, J. & Crow, T. J. 1990 Codon 129 changes in the prion protein gene in Caucasians. *Am. J. Hum. Genet.* **46**, 1215–1216.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. 1992 *Numerical recipes in C—the art of scientific computing*. Cambridge University Press.
- Ramberg, J. S., Tadikamalla, P. R., Dudewicz, E. J. & Myktyka, E. F. 1979 A probability distribution and its uses in fitting data. *Technometrics* **21**, 201–209.
- The National CJD Surveillance Unit 2000 Creutzfeldt–Jakob disease surveillance in the UK. Edinburgh: Western General Hospital.
- Valleron, A. J., Boelle, P. Y., Will, R. & Cesbron, J. Y. 2001 Estimation of epidemic size and incubation time based on age characteristics of vCJD in the United Kingdom. *Science* **294**, 1726–1728.
- Wadsworth, J. D. F., Joiner, S., Hill, A. F., Campbell, T. A., Desbruslais, M., Luthert, P. J. & Collinge, J. 2001 Tissue distribution of protease resistant prion protein in variant Creutzfeldt–Jakob disease using a highly sensitive immunoblotting assay. *Lancet* **358**, 171–180.
- Wells, G. A. H., Dawson, M., Hawkins, S. A. C., Green, R. B., Dexter, I., Francis, M. E., Simmons, M. M., Austin, A. R. & Horigan, H. W. 1994 Infectivity in the ileum of cattle challenged orally with bovine spongiform encephalopathy. *Vet. Rec.* **135**, 40–41.
- Wells, G. A. H., Hawkins, S. A. C., Green, R. B., Austin, A. R., Dexter, I., Spencer, Y. I., Chaplin, M. J., Stack, M. J. & Dawson, M. 1998 Preliminary observations on the pathogenesis of experimental bovine spongiform encephalopathy (BSE): an update. *Vet. Rec.* **142**, 103–106.
- Wilesmith, J. W., Ryan, J. B. M. & Atkinson, M. J. 1991 Bovine spongiform encephalopathy: epidemiological studies on the origin. *Vet. Rec.* **128**, 199–203.
- Will, R. G., Ironside, J. W., Zeidler, M., Cousens, S. N., Estibeiro, K., Alperovitch, A., Poser, S., Pocchiari, M., Hofman, A. & Smith, P. G. 1996 A new variant of Creutzfeldt–Jakob disease in the UK. *Lancet* **347**, 921–925.
- Will, R. G., Zeidler, M., Stewart, G. E., Maclean, M. A., Ironside, J. W., Cousens, S. N., Mackenzie, J., Estibeiro, K., Green, A. J. E. & Knight, R. S. G. 2000 Diagnosis of new variant Creutzfeldt–Jakob disease. *Ann. Neurol.* **47**, 575–582.
- Windl, O. (and 11 others) 1996 Genetic basis of Creutzfeldt–Jakob disease in the United Kingdom: a systematic analysis of predisposing mutations and allelic variation in the *PRNP* gene. *Hum. Genet.* **98**, 259–264.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.