When did bovine spongiform encephalopathy (BSE) start? Implications on the prediction of a new variant of Creutzfeldt-Jakob disease (nvCJD) epidemic

Carine H Cohen and Alain-Jacques Valleron

Background

Knowing the starting date of the BSE epidemic and its size at the very beginning is crucial to interpret the timing of the nvCJD cases and to forecast the nvCJD epidemic. The first cases occurred in 1985. The models devised by Anderson (back-calculation) and Dealler (age-period-cohort) led to an estimate of less than 50 cases in 1983, and none earlier. Here, we applied age-cohort models to the BSE data in order to estimate the earliest possible date of the first unrecognized BSE cases.

Methods

The numbers of confirmed BSE cases in the UK, by age group and by calendar year from 1988 to 1996, were analysed by Poisson regression. The cases’ age distribution was considered as constant between the different birth cohorts. The herd’s age structure was taken into account.

Results

According to the models, BSE cases may have occurred as early as 1980. The expected number of cases before 1990 is almost twice the number of confirmed cases and exceeds by more than 20% the expected value of Anderson’s model. The scenario of first human exposure in 1980 leads to fewer future nvCJD cases than predicted by Cousens with exposure patterns starting in 1983 or 1985.

Conclusion

The first birth cohort available, consisting of two cases older than 10 in 1988, does not allow any projections before 1980. Moreover, confidence intervals are wide and the power of the study is limited by the great dispersion of the data; the precision of the estimations would be improved by considering geographical incidence. Nevertheless, our projections are consistent with Wilesmith’s survey of rendering plants relating the emergence of BSE to the dramatic fall in the proportion of meat and bone meal following solvent extraction, initiated in the late 1970s (65% in 1977 to 10% in 1983).

Keywords

Bovine spongiform encephalopathy, new variant Creutzfeldt-Jakob Disease, age-cohort analysis

Accepted

12 November 1998

The forecast of the epidemic of new variant Creutzfeldt-Jakob disease (nvCJD) relies on two unknown key parameters: incubation time distribution in humans and past exposure to BSE-infected cows. Human exposure was maximum before 1990 (before the ban of specified bovine offals from human food in November 1989 in England and Wales, and in February 1990 in Scotland). The model of Cousens has shown the critical sensitivity of the predictions to the starting point of human exposure. Cousens examined the dates of 1985 (Exposure Pattern A: number of confirmed BSE cases) and 1983 (EPB: number of BSE cases predicted by Anderson’s model using maximum bounds on underreporting before July 1988, but considering no underreporting after July 1988). For instance, given a hypothetic human incubation time of mean 25 years, his prediction of 80 000 nvCJD under the assumption of no infections before 1985 (EPA) turns into 24 000 nvCJD (70% fall) if infections could occur as early as 1983 (EPB).

In fact, the available information is that BSE was identified in 1986 in the UK, and made notifiable in July 1988 (50%
compensation rate in 1988, 100% in 1990). But underreporting was considerable at that time, and will affect any estimation of the starting date. Moreover, the epidemic shows a significant clustering of cases at both geographical and herd levels\(^2\)\(^\text{--}^6\) (variance greater than the mean); this overdispersion phenomenon will also affect any estimation. Such a pattern results from the infection process via cattle food; Wilesmith\(^7\) has related the emergence of BSE to the dramatic fall in the proportion of meat and bone meal, initiated in 1977. As the latency period shows a lag of 2 years, the scenario of first unrecognized BSE cases occurring as early as 1980 cannot be rejected, nor confirmed, on the basis of that survey.

So far, only two models have provided an estimation of the first unrecognized BSE cases. Assuming a mean value of 5 years for the incubation time of BSE, the pre-July 1988 cases have been estimated by Anderson through a back-calculation model.\(^2\)\(^,\)\(^3\) His model assumed no underreporting after July 1988, or after 1990 in univariate sensitivity analyses; as it was designed to make predictions for 1997--2001, and the reporting bias concerning the data in the early phases of the epidemic would not affect the predictions for the future. However, Anderson predicted no cases before 1983,\(^2\) and discussed the limits of this approach. Another model was proposed in 1995 by Dealler: an age-period-cohort model of BSE cases from 1989 to 1993.\(^8\) His model relied on various assumptions about underreporting, and ignored the population’s age structure. He predicted no cases before 1982, and only one in 1982.

Here, to illustrate the plausible importance of underreporting in the early phases of the epidemic, we applied an age-cohort model to the BSE data from 1988 to 1996. Our purpose was to estimate the earliest possible date of the first BSE cases. To achieve that aim, we assumed that the age distribution of the BSE cases remained stable during the whole course of the epidemic. We also took into account the herd’s demography: most infected cows are culled before clinical onset (censoring observation of BSE cases).

## Methods

### Data in cows

The numbers of confirmed BSE cases in Great Britain, documented by age group and year of observation from 1988 to 1996, were derived from parliamentary question 1019 on 10 March 1997 to the Ministry of Agriculture, Fisheries and Food (MAFF). The data were posted on the web. The analyses were conducted on 160 370 cases falling into nine calendar years and ten age groups. The oldest age group consisted of cases aged 11 or older (range: two cases in 1988; 78 cases in 1995). Two one-year-old cases and 5323 cases of unknown age (on average 3% of the cases; up to 8% in 1988) were not considered.

MAFF’s data show that more than 95% of the BSE cases are dairy cows. The dairy herd’s age structure has been estimated in March and December 1996 in England and Wales for 11 age groups, from the cohort born in 1993 (aged two) to the cohort born in 1983 (aged 12). The survival depends on age as a consequence of culling: on average, 27% of dairy cows are culled each year (range: 13--48%).

Considering the culling that followed the March 1996 crisis, we defined the herd at risk for 1996 as the December 1996 dairy herd, and the herd at risk for 1995 as the March 1996 dairy herd. From 1988 to March 1996, few changes took place in the dairy herd’s age distribution and the dairy herd size decreased by 14%. Thus, the age structure of the herd at risk for the years 1988--1995 was taken as the March 1996 dairy herd’s age structure. And for the years 1988–1994, the number of animals at risk in each age group was corrected by the ratio of the dairy herd’s size in 1995 to the dairy herd’s size in each of those years. The oldest age group was made up of 11-year-old and 12-year-old dairy cows.

### Reconstruction of the pre-1990 BSE cases with age-cohort models

Age-period-cohort models assume that the effects of the different factors on the probability of sickness can be separated.\(^9\)\(^,\)\(^10\) For example, age-cohort models were used to analyse the mesothelioma mortality and predict the future course of the epidemic.\(^11\)\(^--^13\) The cohort factor accounts for influences which affect incidence rates in a specified generation or birth cohort equally throughout life. The period factor represents a sudden change affecting equally the individuals of any age simultaneously, typically a modification in reporting rates. The age factor is the biological age effect on incidence, stripped of influences of periods and cohorts. Because the three factors are not independent, the models are beset by problems of identifiability: some solutions have been proposed by applying constraints, but they remain arbitrary and controversial.\(^9\)\(^,\)\(^10\) Such an approach was adopted by Dealler\(^8\) to analyse the BSE cases aged 2--10, from 1989 to 1993: he assumed that there was no underreporting in 1990 and 1991, and that underreporting in 1992 and 1993 was different for cases born in 1988 (cattle food ban) or after. In the models with two factors, the identifiability problems are solved by choosing a parameterization, so that the parameters look like age-specific rates for the age factor, and have direct interpretations in terms of relative risks for the period or for the cohort factor, taking one period or one cohort as reference.\(^9\)\(^,\)\(^10\)

Given that the cow was not culled before showing BSE, the age of a BSE case is the sum of the age at infection (which depends on exposure levels and susceptibility as a function of age) and the incubation time. Assuming that no changes occurred in age at infection and in incubation time across periods or cohorts, the age factor can be estimated either by age-cohort models, or by age-period models, which predict constant ratios of age-specific rates between different cohorts, or different periods, respectively.

The probability of infection depends on the exposure level to the prion in cattle food, and the cows were often infected in their first year of life.\(^2\)\(^,\)\(^4\)\(^,\)\(^6\)\(^,\)\(^14\) Because of the long average latency (5 years), and its variability between subjects, a change in population exposure is more likely to manifest several years subsequently, and not simultaneously in all age groups. If certain generations had greater exposure than others, these changes would lead to a cohort factor. As the variations in exposure levels were wide\(^2\)\(^,\)\(^4\)\(^,\)\(^6\) (the rise due to the recycling of the infectious agent through cattle food was followed by a decrease after the food bans of July 1988 and September 1990), the temporal variations of BSE incidence rates should better be interpreted in terms of the birth cohort to whom the rates apply, with age-cohort models, rather than in terms of the calendar period of report, with age-period models.
The distribution of cases by age group and by birth cohort was analysed by Poisson regression, fitting the age-cohort multiplicative model:

$$ R_{ij} = e^{K \cdot e^{A_i} \cdot e^{C_j}} $$

where $R_{ij}$ denotes the expected incidence rate in age group $i$ and birth cohort $j$, $K$ is a constant, $A_i$ the effect of age group $i$ ($i = 1, \ldots, 10$), and $C_j$ the effect of birth cohort $j$ ($j = 1, \ldots, 18$). The incidence rates were computed as the ratio of the number of cases (range: four cases aged two in 1996; 15 181 cases aged four in 1992), to the corresponding number of cows at risk (range: 18 575 cows aged 11 or 12 in 1996; 664 879 cows aged two in 1988). To investigate the influence of the herd’s demography, the size of the herd at risk was set as a constant (an ‘offset’), and the binomial distribution (large population, low incidence) was approximated by the Poisson distribution. The model’s parameters were estimated by the maximum likelihood method, using the GLIM© software. The estimated values of these parameters remain identical if an overdispersion parameter (which accounts for other effects than those of the factors) is estimated, but confidence intervals become wider.

**Prediction of nvCJD cases adopting Cousens’ model**

The model developed by Cousens was aimed to forecast the size of the future epidemic of nvCJD in the UK, under various hypotheses about the distribution of incubation time in humans (he examined several distributions). His predictions were based on past yearly numbers of BSE cases, used as an indicator to quantify the past human exposure to the BSE-infected cows that entered the human food chain.

Since 1990, a 10-fold factor was assumed to account for the efficiency of measures undertaken on human food: the numbers of BSE cases divided by 10 were used as the exposure pattern. Before 1990, Cousens considered two exposure patterns: one starting in 1985 (EPA) and one starting in 1983 (EPB).

Here, we compared the predictions obtained with the exposure pattern which starts in 1985 (EPA), to those obtained with the age-cohort model’s estimations of BSE cases (called ACEP), starting in 1980.

**Results**

The sensitivity analyses have shown that the rough approximations concerning the figures of the herd at risk (see data section in Methods) would not modify the figures of the estimates presented hereafter.

Figure 1a displays the age factor and Figure 1b displays the cohort factor, in arbitrary units. Due to the herd’s age structure, the reduction in the risk (incidence) in old ages (Figure 1a) is less important than the reduction in the number of cases observed. Such a risk, almost increasing in old ages, may also reflect a part of the herd that was exposed at an older age (various feeding practices, whether calves were exposed to infected food), or reflect genetically distinct populations with different latency periods. The model did not handle satisfactorily the overdispersion of the data; CI were wide. However, the figures we found were really close to those of the back-calculation model of Anderson. For instance, we estimated the age factor’s mean value at 6.35 years (Figure 1a). This may be compared to the Anderson’s estimate of 1.31 years for the average age at infection, assuming a mean incubation time of 5 years and considering that incubation time does not depend on age at infection: $1.31 + 5 = 6.31$.

Table 1 shows that the expected total number of cases before 1990 is almost twice the number of confirmed cases and exceeds the expected value of Anderson’s model by more than 20%. In 1989 for instance, 8660 cases are predicted while 7823 cases were reported, which represents a difference of 837 cases. The first BSE cases predicted by the age-cohort model occurred in 1980 (ACEP), which is 3 years earlier than predicted by Anderson (EPB) and 5 years earlier than the first reported cases (EPA).

Figure 2 displays the yearly timing of two extreme human exposure patterns, EPA and ACEP, on the scale of the numbers of BSE cases divided by 100; and an example of the yearly timing of the nvCJD that may result from these exposure patterns, assuming a hypothetic incubation period of mean 25 years.

Table 2 gives an example of the numbers of nvCJD predicted by adopting Cousens’ model for two exposure patterns given in Table 1: the confirmed BSE cases (EPA) and age-cohort model’s predictions (ACEP). In this example, an epidemic of 1546 nvCJD with EPA with a plausible scenario achieved with a 25 years mean human latency period and displayed by Figure 2) is limited to 1124 nvCJD with the pre-1990 exposure pattern provided by the age-cohort model (ACEP); this represents a 27% fall. This reduction may be only 10%, from 183 to 165 cases, in the optimistic hypothesis of a 10-year mean latency.
Table 1 From 1980 to 1989, numbers of BSE cases (ACEP) estimated by the age-cohort model applied to BSE incidence rates. The population at risk was taken as the dairy herd in the UK (December figures). For comparison, reproduced from Dealler's web site⁸ and from Cousens' article¹.

<table>
<thead>
<tr>
<th>Period</th>
<th>Dairy herd in the UK</th>
<th>Confirmed BSE cases (EPA)⁸</th>
<th>Anderson’s predictions (EPB)³</th>
<th>Dealer’s predictions⁴</th>
<th>AC model’s best estimates (ACEP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td>3,277,000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>1981</td>
<td>3,293,000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>83</td>
</tr>
<tr>
<td>1982</td>
<td>3,353,000</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>272</td>
</tr>
<tr>
<td>1983</td>
<td>3,429,000</td>
<td>0</td>
<td>25</td>
<td>41</td>
<td>536</td>
</tr>
<tr>
<td>1984</td>
<td>3,311,000</td>
<td>0</td>
<td>50</td>
<td>276</td>
<td>818</td>
</tr>
<tr>
<td>1985</td>
<td>3,256,000</td>
<td>63</td>
<td>240</td>
<td>767</td>
<td>1,188</td>
</tr>
<tr>
<td>1986</td>
<td>3,242,000</td>
<td>63</td>
<td>1,052</td>
<td>1,526</td>
<td>1,849</td>
</tr>
<tr>
<td>1987</td>
<td>3,052,000</td>
<td>662</td>
<td>3,310</td>
<td>2,851</td>
<td>3,058</td>
</tr>
<tr>
<td>1988</td>
<td>2,975,000</td>
<td>3238</td>
<td>4,760</td>
<td>5,078</td>
<td>5,260</td>
</tr>
<tr>
<td>1989</td>
<td>2,932,000</td>
<td>7,823</td>
<td>8,203</td>
<td>8,660</td>
<td>8,660</td>
</tr>
<tr>
<td>Total 1980–1989</td>
<td>11,801</td>
<td>17,260</td>
<td>18,743</td>
<td>21,737</td>
<td></td>
</tr>
</tbody>
</table>

a Numbers of confirmed BSE cases (EPA).

b Numbers of BSE cases predicted by Anderson’s back-calculation model (EPB), using maximum bounds on underreporting before July 1988, but considering no underreporting after July 1988.

c Numbers of BSE cases predicted by Dealler’s age-period-cohort model, taking into account underreporting, but assuming no cattle born after 1991 would ever develop BSE.

These projections must be considered in view of the cases predicted before 1994. Cousens explained that it is likely that very few nvCJD could have been missed in young people at that time. Therefore, the main result is that, whatever the plausible latency time distribution assumed in humans, the exposure pattern called EPB, and even to a larger extend ACEP, do not allow for huge epidemics such as those that could be predicted under EPA, because they lead to too many pre-1994 nvCJD. In sensitivity analyses, the nvCJD forecasts depend more on the starting date of exposure, than on exact yearly numbers of BSE.
Table 2 Numbers of nvCJD cases predicted, using an hypothetic incubation time distribution in humans (mean of 10 years or 25 years), so that exactly 23 cases are expected in the years 1995–1997. Since 1990, the human exposure pattern was taken as the numbers of confirmed BSE cases divided by 10. The pre-1990 exposure pattern was taken as the numbers of BSE cases (EPA), that started in 1985; or as the numbers of BSE cases predicted by the age-cohort model (ACEP), that started in 1980 (Table 1). For comparison, reproduced from Cousens’ article

<table>
<thead>
<tr>
<th>Starting of human exposure (Exposure Pattern)</th>
<th>1985 (EPA)</th>
<th>1983 (EPB)</th>
<th>1980 (ACEP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean incubation time</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25 years</td>
<td>80 000</td>
<td>24 000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>13 000</td>
<td>7000</td>
<td></td>
</tr>
<tr>
<td>10 years</td>
<td>213</td>
<td>151</td>
<td></td>
</tr>
<tr>
<td></td>
<td>211</td>
<td>156</td>
<td></td>
</tr>
</tbody>
</table>

* Numbers of nvCJD predicted for different incubation time distributions.

The pre-1990 exposure pattern was taken as the numbers of BSE cases (EPA), that started in 1985; or as the numbers of BSE cases predicted by Anderson’s back-calculation model (EPB), that started in 1983.

**Discussion**

The choice of age-cohort models was determined by the fit to the data, which was worse with age-period models. Nevertheless, the possibility of a period factor due to changes in registration completeness is still not ruled out. The striking rise in cases born after the food bans of 1988 and 1990, and the changes in compensation arrangements for BSE cases (in January 1994, this became based on the price of an older cow instead of the average price for one) may have led to some underreporting, at least before March 1996, when the crisis has led to a shortened reporting delay. And at the beginning of the epidemic, underreporting was certainly important, when notification and increase in compensation rate (50% in August 1988 to 100% in February 1990) may have led to some underreporting, at least before March 1996, when the crisis has led to a shortened reporting delay.2 And at the beginning of the epidemic, underreporting was certainly important, when notification and increase in compensation rate (50% in August 1988 to 100% in February 1990) have increased declaration rates.18 Indeed, for the earliest cohorts, the lack of fit of the age-cohort model can be interpreted in terms of underreporting.

If so, it may be surprising that so many BSE cases were missed by farmers and veterinary surgeons. In 1989, could 837 cases (8660 predicted minus 7823 reported) of BSE, a newly notifiable disease, have been underreported? These should be compared to the dairy herd of 2.9 × 10^6 cows in 1989, keeping in mind that the first symptoms are not specific and may be confused with other nervous disorders in cattle.

Projections based on only two cases older than 10 in 1988 (first cohort available) are subject to great uncertainty. Furthermore, the key hypothesis of a constant age distribution of the BSE cases implies that the age at infection distribution and the incubation period distribution are similar among the different birth cohorts; these assumptions have already been discussed by Anderson.2,3 This hypothesis can lead to overestimate BSE numbers in the 1980s: it is plausible that the first infections occurred at an older age, in the case that the exposure to contaminated food was reduced or occurred later. But the projections rely on the BSE cases observed in the late 1980s, when underreporting was important. Therefore, there may have been much more than two cases older than 10 in the 1977 birth cohort (Dealler argued that reporting rates as low as 0.39 may account for the progressive compliance to legislation); and this would bias the estimations in the opposite direction. As BSE may result from the agent of scrapie in sheep, and incubation is longer in the first passage of these agents to another species, the latency may have been longer in the earliest cohorts, and the age at infection not necessarily older. Finally, the main point is that our projections are consistent with the only epidemiological enquiry conducted at that time, which has shown that the changes in the rendering process were initiated in 1977.

The first birth cohort available, consisting of two cases older than 10 in 1988, does not allow any projections before 1980, even with models considering underreporting in the first periods, or a different latency in the first cohorts. Thus, given the extreme assumption of a constant age distribution of BSE cases, the results we achieved should be interpreted as estimates of past underreporting (upper-bound estimates if the first cohorts were infected at an older age), and the date of 1980 may represent the earliest possible date of the first unrecognized BSE cases. This date leads to more optimistic scenarios concerning the future nvCJD epidemic in the UK than the dates of 1983 or 1985.

**Conclusion**

As Anderson explained, more insight in the early phases of the BSE epidemic would require detailed models. Here, we used a very simple model to describe the BSE epidemic. This model led to estimate the earliest possible date of human exposure in 1980, in agreement with Wilesmith’s epidemiological study. Due to the rough quality of the data, the fit was poor (wide CI). The over-dispersion problem has been discussed elsewhere. Concerning BSE, more detailed data are required for age-cohort models to investigate the clustering of cases in time (quantified by the BSE cases). From a pragmatic viewpoint, despite data pitfalls and epidemiological uncertainties, the optimistic consequences (less nvCJD expected) of an earlier exposure should be considered in public health risk assessment, to provide a rationale in political decisions. Moreover, from the public health standpoint, it is important to emphasize that a starting date of exposure in 1980, while the first human deaths occurred in 1995, is consistent with the assumption of a long incubation time in humans (with a long lag period), and with the hypotheses that most of the nvCJD may have been infected at a young age and that the susceptibility in humans may depend on age.
References


