European Journal of Epidemiology (2006) 21: 47–54
DOI 10.1007/s10654-005-4638-z

NEUROLOGICAL DISEASES

Evaluation of the cumulative evidence for freedom from BSE in birth cohorts

Dankmar Böhning[1] & Matthias Greiner[2,3]
[1]*Working Group in Biostatistical Modelling and International Epidemiology, Section of Applied Statistics, School of Biological Sciences, The University of Reading, Earley Gate Whiteknights Road, Reading, RG6 6FN, UK;* [2]*International EpiLab, Danish Institute for Food and Veterinary Research, Mørkhøj Bygade 19, DK-2860, Søborg, Denmark;* [3]*OIE Collaborating Centre for Research and Training in Population Animal Health Diagnosis and Surveillance Systems*

Accepted in revised form 1 November 2005

Abstract. Substantial resources are used for surveillance of bovine spongiform encephalopathy (BSE) despite an extremely low detection rate, especially in healthy slaughtered cattle. We have developed a method based on the geometric waiting time distribution to establish and update the statistical evidence for BSE-freedom for defined birth cohorts using continued surveillance data. The results suggest that currently (data included till September 2004) a birth cohort of Danish cattle born after March 1999 is free from BSE with probability (power) of 0.8746 or 0.8509, depending on the choice of a model for the diagnostic sensitivity. These results apply to an assumed design prevalence of 1 in 10,000 and account for prevalence heterogeneity. The age-dependent, diagnostic sensitivity for the detection of BSE has been identified as major determinant of the power. The incorporation of heterogeneity was deemed adequate on scientific grounds and led to improved power values. We propose our model as a decision tool for possible future modification of the BSE surveillance and discuss public health and international trade implications.

Key words: Birth cohorts, Bovine spongiform encephalopathy, Freedom from disease, Geometric waiting time, Power, Prevalence heterogeneity

Introduction

In the European Union, a large number of cattle are tested annually for bovine spongiform encephalopathy (BSE). In 2003, a total number of 10.041.295 cattle was tested in the 15 member states [1]. The test costs vary between €43 and €90 (in the UK up to €350) in the member states and the community subsidises the testing with €15–30 per animal tested, depending on the target group [2]. Animals eligible for BSE testing fall into one of the following categories, sometimes referred to as 'surveillance streams': fallen stock (FS), clinical suspects (CS), emergency slaughter (ES) and healthy slaughtered (HS) cattle. Moreover, cattle culled in connection to a BSE case are subject to BSE testing. Only the FS, CS and ES categories are considered as 'risk animals' [1]. The age limit to test HS cattle differs between countries. The EU regulation sets this limit at 30 months but Germany, France, Spain and Italy have chosen a 24 months age limit. The BSE detection rates are generally low, especially in HS animals. In Denmark, one BSE case was found in the HS stream in 2002 and in 2003. This corresponds to costs per case detected of €10,950,720 in 2002. Alternative surveillance schemes are therefore of interest that are economically more sustainable without compromising the protection of consumers. BSE cases are not homogeneously spread over the birth cohorts. This has been shown for Portugal [3], the UK [4], France [5] Switzerland [6], Germany [7], Spain and Ireland [8] and Denmark [9] and can be explained by a higher risk of exposure in the past. Therefore, we use the biological assumption that young cattle may be free of BSE while elder cattle of the same population may be infected at low prevalence due to exposure in the past. Furthermore, we stipulate that the occurrence of infection constrained to distinct birth cohorts may constitute an epidemiological barrier that would comply with the principle of 'compartmentalization' endorsed by the World Animal Health Organisation (OIE). With this motivation, we were interested to develop a method for testing distinct (younger) birth cohorts of cattle free of BSE. The power function for documenting freedom from BSE should account for diagnostic misclassification and prevalence heterogeneity. The issue of optimal sample allocation should be addressed. The methods should be used with Danish BSE surveillance data as case study.

DB conducted this research at the International EpiLab in Copenhagen, Denmark.

Materials and methods

Derivation of the power function

Consider a continuous stream of BSE surveillance data obtained from individuals of a population of cattle. Let the binary variable Y_t denote the test result ('trial') for animal t, where $y_t = 1$ denotes that the animal is positive and $y_t = 0$ otherwise. Interest is in the null hypothesis $H_0{:}\pi = 0$, where π is the prevalence parameter of interest. Clearly, the probability $P(Y_t = 1|H_0)$ is zero for all times $t = 1,2,\ldots$, e.g., the probability of a type-I error is zero. Suppose now as alternative hypothesis that there is some positive (potentially small) prevalence $\pi > 0$. An appropriate model for discrete waiting time T is the geometric distribution $P(T = t|\pi) = (1 - \pi)^{(t-1)}\pi$ for $t = 1, 2,\ldots$, given that $\pi > 0$. The waiting time refers to the number of trials t needed to give the first positive outcome. We are now interested in the smallest stopping time s such that we achieve a specified power $(1 - \beta)$. In our setting, the power is given by $\varphi(\pi, s) = P(0 < T \leq s|\pi > 0) = \Sigma_{t=1}^{s}(1 - \pi)^{(t-1)}\pi$, which can be simplified to

$$\varphi(\pi, s) = 1 - (1 - \pi)^s. \tag{1}$$

In the situation encountered, the choice of a type-II error of $\beta = 0.05$ or below will be appropriate. A positive trial before reaching s provides clear evidence for the cohort being not disease-free and the surveillance will be continued only for reasons of consumer's health protection. If the stopping time s is reached without encountering a positive trial, it is concluded that the cohort is free at a power level of at least $(1 - \beta)$. Equating (1) with the desired power yields the stopping time

$$s = \left\lceil \frac{\log(\beta)}{\log(1 - \pi)} \right\rceil. \tag{2}$$

The number s provides the number of cattle tested to reach a power of $(1 - \beta)$. In addition, one can establish the design prevalence for a given stopping time and power as

$$\varphi^{-1}((1 - \beta)|s) = 1 - \sqrt[s]{\beta}, \tag{3}$$

the smallest design prevalence, which will give a power of $(1 - \beta)$ with s cattle tested negative.

Incorporating misclassification

Here we are concerned with false negative classifications, i.e., lack of diagnostic sensitivity (α). We presuppose that $\alpha < 1$, while the diagnostic specificity is 1. Thus, the unconditional probability for the diagnostic test to deliver a positive result is $\alpha\pi < \pi$. In fact, the most influential factor for the sensitivity is the stage of infection, for which we can only use age as a proxy. Infected animals younger than 24 or 36 months of age are most likely incubating and will test negative. Let $a = 1,\ldots, A$ denote the age class in years. Then, $P(T_a > s_a > 0|\alpha_a, \pi > 0) = (1 - \alpha_a\pi)^{s_a}$ denotes the likelihood for the event that the waiting time T_a for the first animal from the subpopulation of cattle aged a years testing positive is above s_a. This means, $(1 - \alpha_a\pi)^{s_a}$ denotes the probability for a type-II error of incorrectly assigning the cohort of cattle as negative based on all infected cattle with age a having a negative test outcome. Given a vector of positive, integer stopping in times $\mathbf{s} = (s_1, s_2,\ldots, s_A)'$, we are then interested in the event that there exists an age-group a such that the waiting time $T_a \leq s_a$, since in this us case the trial would be stopped. Note that inference is made on the basis of all cattle of the birth cohort, regardless of the age at testing. Now, given the prevalence π and a vector of age-specific sensitivities $\alpha = (\alpha_1, \ldots, \alpha_A)'$ we have that the power is provided as the probability that there exists at least one age group a such that the waiting time $T_a \leqslant s_a$, given the sensitivity and design prevalence, or briefly

$$\begin{aligned} \varphi(\pi, \alpha, \mathbf{s}) &= 1 - P[(T_1, T_2, \ldots, T_A)' \\ &\qquad > (s_1, s_2, \ldots, s_A)'|\alpha, \pi > 0] \\ &= 1 - \prod_{a=1}^{A}(1 - \alpha_a\pi)^{s_a} \end{aligned} \tag{4}$$

We can show that the power obtained in the misclassification model (4) cannot be greater than the power from the simple model (1), which is consistent with our intuitive appreciation of the situation. Age-specific estimates of sensitivity were derived from models for the age at infection and the incubation time [10]. One might compute the probabilities for disease detectability given infection for discrete ages a as $\int_a^{a+1} f(a')da'$, utilizing the density of the incubation period. Similar computations were done by Stockmarr (unpubl. data, 2004) who readily computed these probabilities in a slightly different fashion $\int_{a-0.5}^{a+0.5} f(a')da'$. to combine the distribution of age-at-infection and incubation time distribution in a basic convolution to yield the distribution of an animal becoming a case. Paisley (unpubl. data, 2004) obtained a discretized approximation of the convolution of the two densities by Monte-Carlo integration (λ_a in Table 1). We denote with $\alpha_a = \Sigma_{a'=2}^{a}\lambda_{a'}$ the likelihood that an animal becomes a case in the interval from a to $a+1$ or before and use this as approximation of the sensitivity (α_a in Table 1). Alternatively, it might be argued that, since the animal has lived disease-free up to age a, the likelihood of disease detectable should be computed conditional upon having survived disease-free at age a as $\alpha'_a = \lambda_a/\Sigma_{a''=a}^{A}\lambda_{a''}$ (α'_a in Table 1). It can be shown that α'_a is lower than α_a for all ages.

Incorporating heterogeneity in the design prevalence

We consider here discrete covariates (risk factors for the expected prevalence), so that it is possible to

Table 1. Likelihood for age-specific time to detectability and associated sensitivities[1]

Age group a[2]	Likelihood λ_a	Sensitivity α_a	Sensitivity α'_a
2	0.0001	0.0001	0.0001
3	0.0113	0.0114	0.0113
4	0.1182	0.1296	0.1196
5	0.2642	0.3938	0.3035
6	0.2467	0.6405	0.4070
7	0.1630	0.8035	0.4534
8	0.0950	0.8985	0.4835
9	0.0552	0.9537	0.5438
10	0.0307	0.9844	0.6631
11	0.0100	0.9944	0.2160
12	0.0039	0.9983	0.6964
13	0.0008	0.9991	0.4706
14	0.0008	0.9999	0.8889
15	0.0001	1.0000	1.0000

[1] Sensitivity estimates α_a and α'_a are based on the cumulative likelihood for detectability without and with accounting for survival up to age a, respectively.
[2] Age group a refers to years in the interval $[a, a+1)$.

summarize them in covariate combinations or risk scores with values $r = 1,\ldots, R$. For each of the subpopulations r, there is an associated prevalence π_r. Let now $P(T_{ar} > s_{ar} > 0|\alpha_a, \pi_r > 0) = (1 - \alpha_a\pi_r)^{s_{ar}}$ denote the likelihood for the event that the waiting time T_{ar} for the first animal from the subpopulation of cattle aged a years and risk score r testing positive is above s_{ar}, where $a = 1,\ldots, A$ and $r = 1,\ldots, R$ denote age groups and risk scores, respectively. Given a matrix of positive, integer stopping times

$$\mathbf{s} = \begin{pmatrix} s_{11} & s_{12} & \ldots & s_{1R} \\ s_{21} & s_{22} & \ldots & s_{2R} \\ \ldots & \ldots & \ldots & \\ s_{A1} & s_{A2} & \ldots & s_{AR} \end{pmatrix},$$

and a similarly defined matrix of waiting times $\mathbf{T}$, we are interested in the event that there exists an age-group a and a risk score r such that the waiting time $T_{ar} \leq s_{ar}$, since in this case the trial would be stopped. Now, given a vector of risk score specific prevalences $\pi = (\pi_1, \ldots, \pi_R)'$ and a vector of age-specific sensitivities $\alpha = (\alpha_1, \ldots, \alpha_A)'$ we have that the power $\phi(\pi, \alpha, \mathbf{s})$ is given as the probability that there exists at least one age group a and a risk score r such that $T_{ar} < s_{ar}$, given the sensitivity and prevalence, which is equivalent to 1 minus the probability that $T_{ar} > s_{ar}$ for all age groups a and all risk score groups r, or briefly

$$\begin{aligned} \varphi(\pi, \alpha, \mathbf{s}) &= 1 - P(\mathbf{T} > \mathbf{s}|\alpha, \pi > 0) \\ &= 1 - \prod_{r=1}^{R}\prod_{a=1}^{A} P(T_{ar} > s_{ar}|\alpha_a, \pi_r > 0) \\ &= 1 - \prod_{r=1}^{R}\prod_{a=1}^{A} (1 - \alpha_a\pi_r)^{s_{ar}}. \end{aligned} \quad (5)$$

We propose to select differential design prevalences for the subpopulations of HS cattle and those belonging to any of the risk groups (ES, FS and CS). The risk ratios reported for Denmark are based on a total of 3 and 2 cases for the years 2002 and 2003, respectively, and are therefore associated with a large statistical uncertainty. Using the data from all 15 EU countries [1, 11], we established the combined Mantel–Haenszel risk ratio as $RR_{\mathrm{MH}} = 18.2$ (95% confidence interval 15.3–21.8) for 2002 and $RR_{\mathrm{MH}} = 15.3$ (13.0–18.0) for 2003. Based on these empirical results, a risk ratio of 15 was chosen. For the power analysis, we specify the differential design prevalences such that the minimum value applies to the low-risk group of cattle (HS) and the inflated design prevalence (by factor 15) applies to the high-risk group of cattle.

Optimal sampling design

Suppose, the total of sampled cattle is fixed at $s = \Sigma_{a,r}\, s_{ar}$. We consider now an optimised sampling scheme s_{ar}, which minimises the probability of detection failure,

$$\prod_{a,r}(1 - \alpha_a\pi_r)^{s \times p_{ar}} \quad \text{subject to} \quad \sum_{a,r} p_{ar} = 1, \quad (6)$$

where the p_{ar} are non-negative real numbers and denote the relative frequencies of the population cells defined by age and risk. Taking logarithms and defining $w_{ar} = \log(1 - \alpha_a\pi_r)$ we see that (6) is equivalent to minimizing $\Sigma_{a,r}\, p_{ar}w_{ar}$ subject to $\Sigma_{a,r}\, p_{ar} = 1$ which can be accomplished by noting that $\Sigma_{a,r}\, p_{ar}w_{ar} \geq (\Sigma_{a,r}\, p_{ar})\, \min_{a,r} w_{ar} = \min_{a,r} w_{ar}$. In other words, the linear function $\Sigma_{a,r} p_{ar}w_{ar}$ is minimized at a corner of the probability simplex, where the corner is determined by the combination (a,r) such that $w_{ar} = \min_{a',r'} w_{a'r'}$.

Data for the case study

Surveillance for transmissible spongiform encephalopathy (TSE) in Denmark follows Chapter A of Annex III in Regulation (EC) No 999/2001 of the European Parliament and of the Council as amended by Commission Regulation (EC) No 1494/2002. All cattle that meet the inclusion criteria for the different surveillance streams are tested. There is no random sampling or any other sample selection procedure in place. Data from January 2001 till September 2004 were retrieved from the Danish TSE-database, which is operated under the auspices of the Danish Veterinary and Food Administration. The project database contained the date of birth and death, submission cause (identifying the surveillance stream) and the BSE test result.

Results

Application of the power function

The power function (1), established under assumption of perfect diagnostic tests, is strictly monotone increasing as a function of s as well as a function of π (Figure 1), which can be proven directly. It follows that any prevalence larger than the chosen design prevalence value will reach or exceed the power for the chosen prevalence. For example, after testing $s = 100{,}000$ animals, we reach a power of 0.99995 and 0.63212 under assumption of a design prevalence of 10^{-4} and 10^{-5}, respectively. The latter prevalence has been recommended elsewhere [12]. Furthermore, we use (2) and find that $s = 460{,}515$ and 690,773 animals would have to be tested for a power level of 0.99 and 0.999, respectively, under the assumption of a design prevalence of 10^{-5}. Finally, we use (3) to find the smallest detectable prevalence such that the trial with the given stopping time s achieves a fixed power $1 - \beta$. For illustration, suppose that $s = 30{,}000$ animals have been tested. A design prevalence of 16 or 24 out of 100,000 can be reached with power of 0.99 and 0.999, respectively.

Application to Danish BSE surveillance data

Demography of the birth cohort and preliminary results
During the study period, 13 BSE cases were observed. The distribution of the birth dates is provided in Figure 2. The last (youngest) case was born in March 1999. In August 2005, a suspect case was found, which was born in 1996. Due to the birth date, this case (if confirmed) is not relevant for our analysis. The birth dates of the Danish BSE cases cluster between 1996 and 1997, though isolated cases were born in 1998 up to early 1999. Therefore, it appears best justified to consider Danish cattle with birth date after March 1999 as a candidate free birth cohort. The number of cattle tested in the candidate free birth cohorts are shown in Table 2. The years 1999 and 2000 contribute almost equally to the cohort. 2001 and 2002, in particular, contribute less, caused by the exclusion of all cattle younger than 24 months. We consider first the assumption of a perfect diagnostic test. Given the observed number of $s = 286{,}742$ animals tested from the candidate free birth cohort, a prevalence of 10^{-5} can be detected with power of 0.94316.

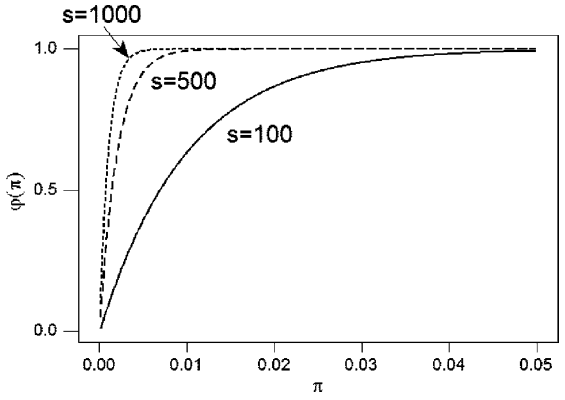

Figure 1. Monotonicity of $\varphi(\pi|s)$ for three different values of $s = 100{,}500{,}1000$.

Power when considering age-specific sensitivity
The candidate free birth cohort includes a high proportion of young animals, which results in a considerably loss of sensitivity (Table 3). More than 80% of all cattle are in the third and fourth age group where sensitivity is still low. Consequently, the power, computed according to (4), is weighted down. For example, we achieve a power of 0.6029 and 0.5693 for a design prevalence of 10^{-4} using the two models for age-specific sensitivities α_a and α'_a, respectively. The corresponding results for the power are 0.1687 and 0.1550 for a design prevalence of 5×10^{-4} and 0.0882 and 0.0808 for a prevalence of 10^{-5}.

Power when considering prevalence heterogeneity
Eighty-seven percent of the Danish birth cohort were healthy slaughtered, whereas 13% can be summarised into the risk categories. The distribution by age group is shown in Table 4. Using these data and a ratio 15 of the design prevalences in risk animals versus HS cattle and under assumption of an age-specific sensitivity, we obtain updated estimates of the power (Table 5). Clearly, the power improves substantially after incorporation of the prevalence heterogeneity, which seems natural and underlines the importance of risk-based components in the surveillance stream.

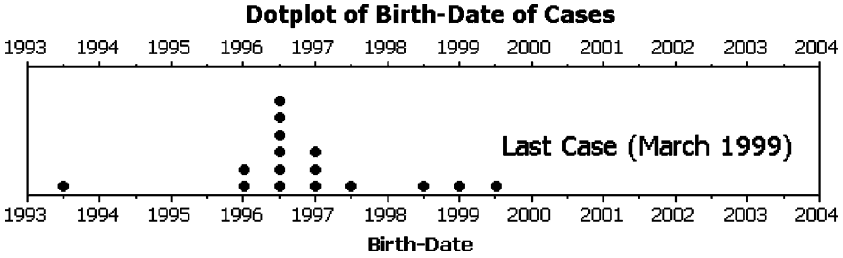

Figure 2. Distribution of birth dates of Danish BSE-cases (confirmed positive) from the Danish TSE-database (January 2001 to August 2004).

Table 2. Distribution of the birth dates in Danish candidate BSE free birth cohort[1]

Month	1999	2000	2001	2002	All
1	0	11,154	5,936	988	18,078
2	0	11,235	5,636	692	17,563
3	0	13,852	6,808	356	21,016
4	17,012	11,285	6,016	152	34,465
5	14,821	9,766	4,744	76	29,407
6	12,748	8,292	3,745	21	24,806
7	14,380	9,131	3,732	11	27,254
8	14,285	9,078	3,167	3	26,533
9	13,397	8,342	2,646	0	24,385
10	12,441	8,112	2,212	0	22,765
11	11,660	7,236	1,791	0	20,687
12	11,654	6,781	1,348	0	19,783
All	122,398	114,264	47,781	2,299	286,742

[1] Status on September 2004.

Optimal sampling design

Result (6) shows that the overall power depends on the relative contribution of each cell of the population, defined by age group a and risk category r. The sampling becomes more efficient when putting more weight on cells with larger contributions to the power. Using the population-cell specific sensitivity and prevalence, we find that the power is maximized by sampling cattle from age groups older than 5 years (for the α_a values) or from the age groups older than 5 years and younger than 10 years (for α'_a values), respectively, and those with the highest design prevalence.

Discussion

Sequential statistical testing approach

The statistical methodology used in this study is derived from the so-called group sequential trial (GST) design. The GST is an experimental design in which groups (or batches) of experimental units are enrolled in a subsequent manner to facilitate interim analyses being conducted. In analogy with those batches, we would recommend using the aggregated data for time intervals of, say 6 months. We have adapted the GST approach for the use with surveillance data and incorporated age-specific misclassification and heterogeneity in the underlying design prevalence for the different surveillance streams. The application to BSE surveillance requires that the target population, i.e., the candidate free birth cohort will be redefined when BSE cases occur in previously free cohorts. The statistical testing with the shrunken data set will result in a loss of power. Obviously, this loss is greatest if a young BSE case is found. The basic principle presented here can be refined by considering a safety margin for the definition of candidate birth cohorts using the birth date of the last case observed. The optimality criterion here is to obtain the largest possible cohort with an acceptable probability that the younger neighbour cohort is not infected (using the prior knowledge of the observed case). Appropriate statistical approaches have been developed in the context of a risk assessment of the safety of culling birth cohorts relative to the index case under new EU regulation [13] that contains provisions for a cohort-culling (Stockmarr and Paisley, unpubl. report, 2005).

Table 3. Age distribution of Danish candidate BSE free birth cohort and associated sensitivities[1]

Age group a[2]	Frequency	Sensitivity α_a	Sensitivity α'_a
2	113,197	0.0001	0.0001
3	119,439	0.0114	0.0113
4	50,888	0.1296	0.1196
5	3,218	0.3938	0.3035

[1] Frequency data from the Danish TSE register, sensitivity estimates from Table 1.
[2] Age group a refers to years in the interval $[a, a+1)$.

Table 4. Frequency of cattle in HS and risk surveillance streams by age group for the Danish candidate BSE free birth cohort

Age group[1]	2	3	4	5	All
HS	90,511	107,692	46,161	3,029	247,393
Risk group	22,686	11,747	4,727	189	39,349
All	113,197	119,439	50,888	3,218	286,742

[1] Age group a refers to years in the interval $[a, a+1)$.

Table 5. Achieved power $\phi(\pi, \alpha, \mathbf{s})$ for the Danish candidate BSE free birth cohort adjusted for sensitivity[1] and for heterogeneity of the design prevalence

Design prevalence[2] (10^{-4})	Adjusted for heterogeneity		Not adjusted for heterogeneity	
	Sensitivity α_a	Sensitivity α'_a	Sensitivity α_a	Sensitivity α'_a
1	0.8746	0.8509	0.6029	0.5693
2	0.6459	0.6139	0.3698	0.3437
3	0.4994	0.4698	0.2650	0.2448
4	0.4049	0.3786	0.2062	0.1899
5	0.3398	0.3166	0.1687	0.1550
6	0.2925	0.2718	0.1427	0.1310
7	0.2566	0.2381	0.1236	0.1134
8	0.2286	0.2117	0.1090	0.0999
9	0.2060	0.1906	0.0975	0.0894
10	0.1875	0.1733	0.0882	0.0808

[1] Age-specific sensitivity estimates α_a, α'_a from Table 1.
[2] The specified design prevalence applies to all surveillance streams when heterogeneity is ignored and to the healthy slaughtered when heterogeneity is accounted for.

Temporal aspects of surveillance

The cumulation of evidence for disease freedom over long periods of time is not a general option. For example, surveillance for highly contagious diseases conducted in the past will not provide useful information about the current disease situation. In BSE, the situation is different because the diagnostic information always refers to exposure and infection in the past. Therefore, data can be summarised and analysed over long periods of time. The life span (survival time) of cattle and thus the demographic representation of birth cohorts in the current standing population accounts in a natural way for any outdating of surveillance information. The surveillance credit points awarded in the OIE surveillance scheme remain valid for 7 years [14]. This limit corresponds to the 95th percentile of the incubation period. In our cohort model, we would rather chose a time limit adjusted to the survival function of cattle. The rationale would be to exclude data from birth cohorts from the analysis that are no longer existent in the standing population. If inference is to be made for distinct birth cohorts rather than pooling all candidate cohorts into one, appropriate stratified analyses can be attempted. However, any stratification will lead to a loss of power due to smaller effective sample sizes.

Statistical model

The important underlying assumption, leading to the geometric waiting time distribution in the cohort model, is the independence of diagnostic testing. This assumption seems reasonable for BSE, because there is no evidence of clustering of infection within Danish herds. Another model assumption is that we are sampling from an infinite population. This results in an underestimation of the power. An adjustment might be necessary for countries with cattle populations too small to reach the required power for the specified candidate BSE free birth cohort.

A BSE surveillance model developed by the Community Reference Laboratory for BSE on request of the European Commission (EC) has now been adopted in the OIE guidelines for BSE surveillance [14]. All factors recognised in this model were also considered in our model: a design prevalence of 10^{-5} of the adult cattle population; a confidence level of 95%; the pathogenesis and pathological and clinical expression of BSE (sensitivity of diagnostic methods used; relative frequency of expression by age; relative frequency of expression within each subpopulation; interval between clinical pathological change and clinical expression); demographics of the cattle population, including age distribution; influence of BSE on culling or attrition of animals from the cattle population via the four subpopulations; percentage of infected animals in the cattle population which are not detected; cattle population numbers stratified by age; the number of cattle tested for BSE stratified by age and by subpopulation. The model defines a target number of credit points, depending on the size of the cattle population and the desired design prevalence (10^{-5} or 5×10^{-4}). A country earns surveillance points through surveillance, whereby animals in different age-risk cells of the population contribute differently to the total score of surveillance points.

Design prevalence

The power is computed using a design prevalence $\pi_{DP} > 0$. Purely formally, prevalences in the interval $0 < \pi \leq \pi_{DP}$ will not be detected with power reached at π_{DP}. Though this procedure (and the underlying problem) is similar in other settings such as when testing an effect size, there is another argument which justifies this approach. Suppose that there is a finite population of 100,000, say. What is the smallest prevalence which can occur? Clearly, it is 1 in 100,000, since 1 animal is the smallest unit to be detected. Any remaining risk is below this detectable risk, and, though it might exist, will not lead to any cases and is therefore a pure theoretical construct. Using a design prevalence, which is small but larger than zero, is therefore well justified.

Modelling sensitivity

The power depends crucially on the assumed sensitivity. Biologically, the sensitivity of the BSE detection denotes the probability that an animal tests positive given it is infected. It should be noted that the definition of 'infection' applies to incubating cattle that show no clinical signs. The diagnostic sensitivity depends on two factors. First, it depends on the analytical sensitivity of the test, i.e., on the probability to detect the characteristic BSE alterations such as detection of pathological prion protein given that such alterations are truly present. This analytical sensitivity is assumed 1 in the cohort model. The postmortem diagnostic tests used in routine BSE surveillance are only partially validated [15, 16] and therefore this value might be an overestimation. Second, the diagnostic sensitivity depends on the probability of these characteristic alterations to be present in infected animals. We used estimates for this probability derived from published models for the age at infection and incubation time, although these models were not fit using data from Denmark. Before presenting the numerical values derived from the model, we have asked experts in the area of BSE surveillance to give us their best estimate of age-specific sensitivity values (results not shown). It turned out that the values of the sensitivity based on α_a and α'_a are in a similar range and rather close in the young age groups. For older age groups, the expert

opinion differs considerably from the range of values provided by α'_a. For completeness, we have re-evaluated the power values reached with quantification of the sensitivity through expert opinion. For example, we found a power achieved with a design prevalence of 10^{-5} of 0.3124. Given sensitivity scenarios like the ones reported by our experts, the surveillance needs to continue for a considerable time to reach acceptable power values. There are no data at hand to finally decide, which values for the sensitivity are more valid for the Danish situation.

Public health aspects

Principles of precaution require that, even in the absence of a final roof, a possible aetiologic link between BSE in cattle and variant Creutzfeldt–Jakob disease in humans is considered [17]. However, it should be noted that the removal of risk materials from the human and animal food chains is the primary risk mitigation measure. Our results can be used as scientific decision basis for a discontinuation of BSE surveillance in HS cattle. According to this adaptive surveillance scheme, no change would occur during the time period before BSE freedom in birth cohorts is declared. After substantiating freedom, all risk animals would still be tested and all risk materials would still be removed (even from HS animals). If BSE cases are found in previously 'BSE-free' birth cohorts, the surveillance should again be extended to cover all cattle including HS. Therefore, we assume that our suggested adaptive surveillance scheme presents no or only negligible additional public health risks compared to the present system.

Trade implications

The use of a distinct birth cohort as target subpopulation for documenting BSE freedom is based on the epidemiological interpretation that BSE risk mainly reflects exposure levels in the past. This interpretation has been given by other authors [18, 19]. It is also reasonable to assume that exposure levels have dropped markedly in 1990, when a ban on feeding protein from ruminants to ruminants has been implemented. This ban, reinforced through European legislation, was successively extended to cover all animal protein sources for feeding ruminants (Janury 1997), a ban of meat and bone meal (March 2000) and processed animal protein as feed stuff for all production animals (January 2001).The concept of using birth cohorts as distinct 'epidemiological' entities is new and may be considered analogue to the OIE-endorsed principle of compartmentalisation, which is also applicable to BSE [14]. Therefore, we suggest that the adaptive BSE surveillance scheme is acceptable in an international trade context because it is base on accepted principles.

Conclusions

This paper describes a novel approach for documenting BSE freedom in cattle populations. Surveillance for BSE is inherently problematic because of the low prevalence, the long-incubation period and the ambiguity of the clinical syndrome. However, the long-incubation period and the well-accepted hypothesis of exposure and infection early in a cow's life present a unique opportunity to accumulate diagnostic evidence about the status of the population over long periods of time. The study has demonstrated that the statistical power for documenting BSE freedom of birth cohorts can be expressed mathematically as a function of the assumed underlying design prevalence, the demographic composition of the cattle tested and the diagnostic sensitivity of the detection methods. The suggested cohort model allows documentation of BSE freedom for distinct birth cohorts by using the surveillance results in a cumulative manner. The BSE situation in Denmark is suitable for application of the cohort model since the youngest case detected was born in March 1999 and younger cattle can be declared as candidate BSE free birth cohort for the documentation of BSE freedom. The age-dependent, diagnostic sensitivity for the detection of BSE is an important part of the cohort model and is mainly a function of the (unknown) stage of infection. Models for the sensitivity introduce uncertainty into the calculations. The cohort model allows different levels of design prevalences for the surveillance streams be assumed (heterogeneity). Presently (data status of September 2004), the statistical power, or confidence for BSE freedom of Danish cattle born later than March 1999, adjusted for heterogeneity and for a design prevalence of 1 in 10,000 is 0.8746 or 0.8509, depending on the choice of a model for the diagnostic sensitivity. Older cattle in risk surveillance streams contribute most to the power on a per-animal-basis. Cattle tested in the HS surveillance contributed importantly (about 30.5%) to the power for the time window of the analysis (from a power partitioning analysis).

It should be considered to switch from testing all HS cattle older than 30 (24) months to a sampling process in the HS stream for birth cohorts that can be classified as BSE free by the cohort model. The sampling procedure should be stratified by age and should put more weight on the older age groups. The cohort model should eventually be modified to account for newer developments in the modelling of the incubation time and the time-at-infection. The interim analysis using the cohort model should be done in regular time intervals. A 6-month interval appears to be appropriate. We also conclude that the adaptive BSE surveillance scheme poses no or only negligible additional public health risk compared to the present

system and is acceptable in the context of international trade.

Acknowledgements

We thank all participants of the project group for advice and support given during the entire study period: Jørgen Nielsen and Mariann Chriél (both Danish Cattle Federation), Anders Stockmarr, Mette M. Andersen, Larry Paisley, Julie Hostrup-Pedersen and Peter Lind (all DFVF), Preben Willeberg and Helene Rugbjerg (both Danish Food and Veterinary Administration). The project was funded by the Direktoratet for FødevareErhverv (DFFE, 3401-65-03-741).

References

1. European Commission. Report on the monitoring and testing of ruminants for the presence of transmissible spongiform encephalopathy (TSE) in the EU in 2003, including the results of the survey of prion protein genotypes in sheep Breeds. Directorate D – Food Safety: production and distribution chain D2 – Biological risks, 2004.
2. Mulder J, Fischler M. Written question E-1821/01 by Jan Mulder (ELDR) to the Commission (2002/C 40 E/112), 2001.
3. Donnelly CA, Santos R, Ramos M, Galo A, Simas JP. BSE in Portugal: anticipating the decline of an epidemic. J Epidemiol Biostat 1999; 4: 277–283.
4. Stevenson MA, Wilesmith JW, Ryan JBM, Morris RS, Lockhart JW, Lin D, Jackson R. Temporal aspects of the epidemic of bovine spongiform encephalopathy in Great Britain: individual animal-associated risk factors for the disease. Vet Rec 2000; 147: 349–354.
5. Morignat EC, Ducrot C, Roy P, Baron T, Vinard JL, Biacabe AG, Madec JY, Bencsik A, Debeer S, Eliazsewicz M, Calavas D. Targeted surveillance to assess the prevalence of BSE in high-risk populations in western France and the associated risk factors. Vet Rec 2002; 151: 73–77.
6. Anonymous. Report on the Assessment of the Geographical BSE-risk of Switzerland. http://europa.eu.int/comm/food/fs/sc/ssc/out136_en.pdf, 2000.
7. Anonymous. Report on the Assessment of the Geographical BSE-risk of Germany. http://europa.eu.int/comm/food/fs/sc/ssc/out120_en.pdf, 2000.
8. Vossen P, Kreysa J, Goll M. Overview of the BSE Risk Assessments of the European Commission's Scientific Steering Committee (SSC) and its TSE/BSE ad hoc Group. http://europa.eu.int/comm/food/fs/sc/ssc/out364_en.pdf, 2003.
9. DFVF. Oversigt over tilfælde af kogalskab (BSE) påvist hos danskfødte kreaturer. http://www.dfvf.dk/Default.asp?ID = 9827, 2004.
10. Ferguson NM, Donnelly CA, Woolhouse MEJ, Anderson RM. The epidemiology of BSE in cattle herds in Great Britain: II. Model construction and analysis of transmission dynamics. Philos Trans R Soc London 1997; 352: 803–838.
11. European Commission. Report on the monitoring and testing of ruminants for the presence of transmissible spongiform encephalopathy (TSE) in 2002. Directorate D – Food Safety: Production and Distribution Chain D2 – Biological Risks, 2003.
12. European Commission. Opinion on Requirements for Statistically Authoritative BSE/TSE Surveys. Adopted by the Scientific Steering Committee (29–30 November 2001). http://europa.eu.int/comm/food/fs/sc/ssc/out238-en.pdf, 2001.
13. EC. Commission Regulation (EC) No 1492/2004 of 23 August 2004 amending Regulation (EC) No 999/2001 of the European Parliament and of the Council as regards Eradication Measures for Transmissible Spongiform Encephalopathies in Bovine, Ovine and Caprine Animals, the Trade and Importation of Semen and Embryos of Ovine and Caprine Animals and specified Risk Material. Official Journal of the European Union, 2004.
14. OIE. Surveillance for bovine spongiform encephalopathy. Appendix 3.8.4 of the International Animal Health Code, 2005.
15. Kubler E, Oesch B, Raeber AJ. Diagnosis of prion diseases. Br Med Bull 2003; 66: 267–279.
16. Biffiger K, Zwald D, Kaufmann L, Briner A, Nayki I, Purro M, Bottcher S, Struckmeyer T, Schaller O, Meyer R, Fatzer R, Zurbriggen A, Stack M, Moser M, Oesch B, Kubler E. Validation of a luminescence immunoassay for the detection of PrP(Sc) in brain homogenate. J Virol Meth 2002; 101: 79–84.
17. Croes EA, van Duijn CM. Variant Creutzfeldt-Jakob disease. Eur J Epidemiol 2003; 18: 473–477.
18. de Koeijer A, Heesterbeek H, Schreuder B, Oberthur R, Wilesmith J, van Roermund H, de Jong M. Quantifying BSE control by calculating the basic reproduction ratio R0 for the infection among cattle. J Math Biol 2004; 48: 1–22.
19. Paisley LG, Hostrup-Pedersen J. A quantitative assessment of the risk of transmission of bovine spongiform encephalopathy by tallow-based calf milk-replacer. Prev Vet Med 2004; 63: 135–149.

Address for correspondence: Dankmar Böhning, Working Group in Biostatistical Modelling and International Epidemiology, Section of Applied Statistics, School of Biological Sciences, The University of Reading, Earley Gate Whiteknights Road, Reading, RG6 6FN, UK
Phone: + 44 (0) 118 378 6211; Fax: + 44 (0) 975 3169
E-mail: d.a.w.bohning@reading.ac.uk