参考資料6－1

July 2006

（原文）

Peer Review of the Estimation of the Ongoing Surveillance Plan for Bovine Spongiform Encephalopathy in the United States

Revised Final Report

Prepared for

U.S. Department of Agriculture
Animal and Plant Health Inspection Service
Riverdale, MD

Contracting Officer's Technical Representative

Chuanfa Guo
U.S. Department of Agriculture
Food Safety and Inspection Service
Washington, DC

Prepared by

Sumeet Patil
RTI International
Health, Social, and Economics Research
Research Triangle Park, NC 27709

RTI Project Number 0208893.024

RTI Project Number
0208893.024

Peer Review of the Estimation of the Ongoing Surveillance Plan for Bovine Spongiform Encephalopathy in the United States

Revised Final Report

July 2006

Prepared for

U.S. Department of Agriculture
Animal and Plant Health Inspection Service
Riverdale, MD

Contracting Officer's Technical Representative

Chuanfa Guo
U.S. Department of Agriculture
Food Safety and Inspection Service
Washington, DC

Prepared by

Sumeet Patil
RTI International
Health, Social, and Economics Research
Research Triangle Park, NC 27709

*RTI International is a trade name of Research Triangle Institute.

Contents

Executive Summary

The Office and Management and Budget (OMB) requires a peer review for important scientific information to ensure the quality of scientific and technical research and guide improvements in the draft before federal agencies disseminate it (OMB, 2004). The Animal and Plant Health Inspection Service (APHIS) is interested in conducting a peer review of their ongoing Bovine Spongiform Encephalopathy (BSE) surveillance plan based on the recent estimate of BSE prevalence in the United States. APHIS requested RTI International's (RTI's) support in conducting a peer review conforming to OMB's guidelines (OMB, 2002; 2004) under RTI's task order contract with the Food Safety and Inspection Service (FSIS).

Specifically, APHIS needs a review of the ongoing surveillance plan to ensure that it meets or exceeds the World Organization for Animal Health (OIE) "type A" surveillance recommendations and guides science-based policy and regulatory decisions on BSE risk mitigation. RTI identified three experts and conducted the peer review according to the statement of work. We present these three reviews in this report in Section 6.

Two of the three reviewers agreed that the sample size estimate of 40,000 samples per annum is sufficient and exceeds the requirement for "type A" surveillance by OIE. They agreed with most of the assumptions APHIS made in estimating the sample plan, except for a few suggestions to improve their plan. These reviewers even suggested possible ways to further reduce the required sample size. For example, Dr. Gardner suggested using a Bayesian approach that considers the effect of a feed ban in the United States whereas Dr. MacDiarmid

suggested changing the proportion of samples from each of the three surveillance streams. To further improve the surveillance plan, Dr. MacDiarmid suggested possible policies at the national level to increase the reporting of clinically suspect animals by farmers and to focus the program more tightly on the animals providing the greater number of "analytic points." Dr. Gardner recommended that APHIS consider the sensitivity of the entire surveillance stream in estimating the required sample size.

The third reviewer, Dr. Morris, opined that the sample size can 'possibly' meet the requirement for "type A" surveillance by OIE. His main concern is that the aging of animals needs to be accurately measured in future surveillance plan and careful consideration should be given to the inaccuracies in estimating the age in the present analysis. He also questioned the accuracy or robustness of APHIS's estimate of 'analytic points' per sample or test. He also identified other aspects of the documentation and analysis that warrant additional clarifications or details. Among above, he recommended emphasis on high risk sub-populations of the cattle based on geographical and risk factors to further strengthen the surveillance plan.

Dr. Morris' review focused on a few additional issues that were out of the scope of the charge to the reviewers. However, these comments are related to previous estimation of BSE prevalence and past surveillance plan data. Given the indirect relevance of these comments and in the spirit of objective review process we have included such beyond-the-charge comments as an annex to his report. We summarize the key points from the annex in the main text of this report to increase the readability.

1 Background and Objective

RTI International (RTI) coordinated external peer review of the ongoing Bovine Spongiform Encephalopathy (BSE) surveillance plan in the United States as requested by the U.S. Department of Agriculture's Food Safety and Inspection Service (USDA, FSIS) and the Animal and Plant Health Inspection Service (APHIS) under this task order. In this report, we present the background information about the peer review, describe the review process, list key questions or the charge to the reviewers, and include the three peer review reports.

APHIS has conducted BSE surveillance in the United States cattle herd with increasing intensity since 1990. Beginning in June 2004, these efforts were enhanced significantly, with a goal of obtaining as many samples as possible from the targeted population in a 12- to 18-month period. Data from these surveillance efforts have been analyzed to estimate the prevalence of BSE in the United States. The conclusions of this analysis have been used to design an ongoing BSE surveillance plan for the United States that sets active surveillance at a level needed to ensure that the prevalence of BSE remains extremely low. This level is designed to meet and exceed the World Organization for Animal Health (OIE) "type A" surveillance recommendations. The surveillance is expected to continue monitoring the BSE status of U.S. cattle as a means to detect any rising BSE prevalence and help guide science-based policy and regulatory decisions on BSE risk mitigation. Therefore, this analysis and the report are scientifically important and deserve an external peer review as per the guidelines by the Office and Management and Budget (OMB) (2004).

The objective of the review is to confirm that the scientific information used in developing the maintenance plan is accurate and complete, the mathematical calculations are used correctly, the conclusions are reasonable given the information presented and the mathematical findings, and the sample size estimate meets the objective of maintenance surveillance. RTI conducted a formal and independent peer review of the ongoing BSE surveillance plan as per the charge prepared by APHIS and conforming to OMB's guidelines for peer review and quality of information (OMB, 2002, 2004).

2 Information on the Ongoing Sample Size Estimation

Active surveillance for BSE was initiated in the United States in 1990 and in 2004 enhanced surveillance program was implemented. Through these efforts, two cases of BSE were identified through March 2006. Based on data collected in the United States over the last 7 years, USDA estimates the prevalence of BSE among U.S. cattle as lower than 1 in a million. Prevalence is expected to decline as long as mitigation efforts that maintain low risk for introduction and spread of the BSE agent among U.S. cattle are maintained.

The principal purposes of ongoing BSE surveillance are to continue to monitor the BSE status of U.S. cattle, to provide mechanisms for detecting rising BSE prevalence among U.S. cattle, and to meet or exceed OIE surveillance recommendations.

USDA estimated that 40,000 samples per year are adequate to meet or exceed “type A” surveillance requirements by OIE to maintain confidence that prevalence is less than one infected animal per million adult cattle.

3 Description of Review Process

RTI conducted the review process in accordance with the OMB guidelines (OMB, 2004). The review process consisted of selecting the reviewers, explaining the scope of the review, facilitating the review, and consolidating the reviews in a single report.

First, we selected three peer reviewers based on their expertise. We initially identified 12 potentially suitable reviewers after discussing the background and objectives of the peer review from FSIS and APHIS. Subsequently, we finalized the list to three reviewers based on their availability and the desired overlap of expertise in the science of BSE, particularly related to animal health; international standards related to animal health; and disease surveillance and supporting statistics. We also considered conflict of interest in the selection process.

Second, we explained the scope of the review in terms of the charge to the reviewers prepared by APHIS. Along with the charge, RTI provided the report on sample size estimate for the BSE ongoing surveillance for the peer review and a draft ongoing surveillance plan as a background document to aid the review. The charge consists of two questions as described in Section 4.

Third, RTI communicated and clarified any questions the reviewers had about the scope of the review or the analysis itself. We communicated the progress and status of the review to APHIS and FSIS regularly and ensured that the reviewers were meeting the objectives of the peer review. We also ensured that the reviewers describe possible ways to address

their concerns instead of only describing the concerns. Subsequently, we communicated APHIS's concerns to the reviewers so they best meet the requirement of the peer review.

Finally, we consolidated the three reviews in this report. We provide brief background information on the review process and include the three peer reviews in Section 6.

To maintain the integrity of the reviews, we present the reviews as separate chapters in this report instead of consolidating the comments by the charge questions. Each reviewer focused on different aspects of the charge questions depending on his area of expertise, and their reporting formats and writing styles also differ. Therefore, reading each review separately can help readers better understand their comments. We have corrected minor typographical errors and reformatted their reports to ensure a minimum level of uniformity of presentation in this report.

4 Charge to the Peer Reviewers

APHIS asked the reviewers to focus their review on the specific questions listed below.

1. Please comment on whether the plan and resulting sample size estimate continued on an annual basis over 7 years is consistent with identifying BSE at 1 infected animal per 1,000,000 adult cattle in the population with a high degree of confidence.

   If you conclude that a smaller sample size would be sufficient to meet our objective, please provide guidance on structuring an adequate sample.

   If you conclude that the study is underpowered with respect to meeting this objective, please provide guidance on a more appropriate sample allocation and sample size.

2. Please discuss the basis of your conclusion, including the justification for any differences in the sampling objective (i.e., other than $1/10^6$) that you would recommend. Given that the goal of the report is to determine the prevalence, examine the agency's conclusions of this analysis. Did the agency use the appropriate models and was the data valid? Focus on the suitability of the methods, the transparency of the approach, and the robustness of the results.

In case of Dr. Morris, we had to further explain that the charge was focused only on ongoing surveillance sample size document and not on any reference or background reports.

5 References

The Office of Management and Budget (OMB). 2002. "Information Quality Guidelines." The Office of Information and Regulatory Affairs, the Office of Management and Budget, the Executive Office of the President, Washington, DC. October 1, 2002.

The Office of Management and Budget (OMB). 2004. "Final Information Quality Bulletin for Peer Review." A Memorandum for Heads of Departments and Agencies. M-05-03. The Office of Management and Budget, the Executive Office of the President, Washington, DC. December 16, 2004.

6 Peer Review Reports

Review of "Sample Size Estimate for BSE Ongoing Surveillance"

by

Stuart C. MacDiarmid, PhD
Principal International Adviser (Risk Analysis), the International Coordination Group of Biosecurity New Zealand, the New Zealand Ministry of Agriculture and Forestry.

Dr. Macdiarmid is currently Principal International Adviser, Risk Analysis, with the International Coordination Group of Biosecurity New Zealand (a department of the New Zealand Ministry of Agriculture and Forestry). His duties involve advising on New Zealand's BSE policies, the BSE policies and programs of trading partners, international standards for the safe trade in animals and animal products, and biosecurity risks and risk analyses. He is also an Adjunct Professor in Veterinary Biosecurity, Massey University and Secretary General of the Terrestrial Animal Health Standards Commission (the 'Code Commission') of the OIE, the World Organization for Animal Health. He has also served as a member of various OIE ad hoc expert groups on BSE, Scrapie and risk analysis. He also chaired the expert group which drafted the OIE's two volume Handbook on Import Risk Analysis. From 2002 to 2005 he was Principal Adviser Zoonoses and Animal Health, New Zealand Food Safety Authority and a member of the Science Group, with responsibilities in the areas of transmissible spongiform encephalopathies and salmonellosis. From 1982 to 2002 he held various technical positions with the New Zealand Ministry of Agriculture. He has authored and co-authored more than 100 scientific papers.

I have been commissioned by Research Triangle Institute to offer critical review of the sample size estimate made by USDA for the ongoing BSE surveillance of the US national cattle herd.

Question: Is the plan and sample size estimate sufficient, if continued on an annual basis over 7 years, to be able to detect BSE at a rate of 1 infected animal per million adult cattle, with a high degree of confidence?

Response: The plan, and sample size, have been derived from recognised statistical techniques (Cannon and Roe 1982) and the peer-reviewed BSurvE model (Wilesmith et al. 2004). The aim of the ongoing surveillance is two-fold; firstly, to detect BSE if present at a rate of 1 per million adult cattle and secondly, to exceed the requirements of the OIE's 'Type A' surveillance (*Terrestrial Animal Health Code*).

As pointed out in the document 'Sample Size Estimate for BSE Ongoing Surveillance', the sensitivity of the BSE surveillance program (that is, its ability to detect BSE at a stated design prevalence) is strongly influenced by the various subpopulations of cattle from which the samples are drawn. Given the perfectly reasonable assumption that the proportion of samples drawn from each of the three surveillance streams (clinical suspects, casualty slaughter and fallen stock) in the next 7 year period will be essentially the same as in the last 7 year period, it is valid to assume that the estimated sample size of 40,000 animals per year will provide a very high degree of confidence that BSE will be detected, if it is present at a rate equal to or greater than 1 per million.

The planned sample size will thus provide assurances that BSE is either not present or is present at a prevalence too low to sustain an epidemic. The planned surveillance will also provide an effective mechanism for detecting an increasing prevalence of BSE in US cattle, in the highly unlikely event that current control measures are inadequate.

Again, on the basis of sampling in the past 7 years, it is evident that the planned sample size will provide surveillance assurances which far exceed the requirements of OIE's 'Type A' surveillance. The document correctly estimates that OIE's 'Type A' surveillance requirements would be met by an annual sample of around 10,500 animals, selected from the three surveillance streams in proportions similar to that sampled in the past 7 years.

Observation and comment: The document 'Sample Size Estimate for BSE Ongoing Surveillance' estimates, on the basis of the past 7 years surveillance and the BSurvE model, that each animal sampled will contribute 9.5 analytic points. I believe this is a reasonable estimate. However, I suggest that USDA could either increase the sensitivity of its ongoing surveillance program, or reduce the number of samples, by working to change the proportion of samples harvested from each of the three surveillance streams.

The analytic points allocated by OIE's *Terrestrial Animal Health Code* are based on the BSurvE model. Article 3.8.4.4. of the Code gives the analytic points for different subpopulations of cattle. Table 2 in that Article shows that a so-called 'clinical suspect', as defined in Article 3.8.4.2., is worth 260 points if it is between 2 and 4 years of age. If between 4 and 7 years of age it is worth 750 surveillance points. Article 3.8.4.2. of the Code defines a so-called clinical suspect as one "…affected by illnesses that are refractory to treatment, and displaying progressive behavioural changes such as excitability, persistent kicking when milked, changes in herd hierarchical status, hesitation at doors, gates and

barriers, as well as those displaying progressive neurological signs without signs of infectious illness …"

In New Zealand in 2005, our organisation mounted a drive to increase the high-value animals being picked up by the BSE surveillance program. We removed the financial disincentives that farmers faced when they called a veterinarian to examine animals meeting the criteria outlined in Article 3.8.4.2. These disincentives included the veterinarian's fees and the cost of disposing of a carcass once the head had been removed (renderers refused to take away carcasses without a head). By insuring that the farmer was not financially penalised for calling the veterinarian to so-called 'clinical suspects' falling in the age range 2 to 7 years, we were able to very dramatically increase the average number of 'analytic points' per sample.

It is possible that such an enhancement could be made to the program planned by USDA and, indeed, the final paragraph of the document 'Sample Size Estimate for BSE Ongoing Surveillance' touches on this. However, I recommend that a greater effort be made to focus the program more tightly on to the animals providing the greatest number of 'analytic points'.

Review of "An Estimate of the Prevalence of BSE in the United States"

by

Ian Gardner, MPVM, Ph.D.
Professor of Epidemiology, School of Veterinary Medicine at the University of California, Davis

Dr. Ian Gardner is a Professor of Epidemiology in the School of Veterinary Medicine at the University of California, Davis. His main expertise is in analytic epidemiology and his research interests include diagnostic test evaluation, risk analysis for livestock diseases and food safety, development of methods for certification of pathogen freedom in animal populations, and the epidemiology and transmission of Johne's disease in cattle. Part of his collaborative research with Dr. Wes Johnson involves application of Bayesian methods to diagnostic testing, prevalence estimation and surveillance problems for animal diseases. He is an author of more than 200 peer-reviewed publications and has served on many national and international committees, panels and review teams.

Executive Summary

The estimate of 40,000 samples per annum is conservative given the fact that prior information about prevalence and the effects of the feed ban have not been incorporated into the sample size estimation procedure. First, this reviewer recommends consideration of a Bayesian approach based on modification of previously-developed methods (references 1 and 2). Such an approach is scientifically justifiable and will yield smaller sample sizes, if prior knowledge about prevalence is modeled. Second, the sensitivity of the entire BSE surveillance system should be estimated and incorporated into the calculations, regardless of whether a frequentist or Bayesian approach is used. Third, sample size calculations for the second purpose for ongoing BSE surveillance should be added.

Specific Tasks

(1) Please comment on whether the plan and resulting sample size estimate continued on an annual basis over 7 years is consistent with identifying BSE at 1 infected animal per 1,000,000 adult cattle in the population with a high degree of confidence.

If you conclude that a smaller sample size would be sufficient to meet our objective, please provide guidance on structuring an adequate sample.

(2) Please discuss the basis of your conclusion, including the justification for any differences in the sampling objective (i.e., other than 1/10^6) that you would recommend.

The selected sampling objective of detection of 1 case in 1 million adult cattle seems reasonable to this reviewer and not warranting change.

The sample size calculations in the document are based on a traditional frequentist approach to the problem. The calculations use the Cannon and Roe formula with adjustments for analytic points and the approach is reasonable, assuming that the sampling plan truly represents the "high-risk" groups. Strictly speaking, the sample size is to detect at least 1 BSE infected animal with high confidence rather than to estimate prevalence. However, if the sampling is random unbiased estimates of prevalence and upper 95% confidence intervals can be obtained. The calculation assumes that 1) the true prevalence in the population is constant over the 7-year period or it can be interpreted as the "average value" for the 7-year period, and that 2) the sensitivity and specificity of the entire BSE surveillance system in the U.S. are both 100%.

I will briefly comment on each of those assumptions:

1. Based on the assumption that the feed ban is an effective mitigation and that older infected cattle will be removed from the population by culling, it is justifiable to believe that prevalence is declining and will continue to decline over time.

2. Although the BSE diagnostic tests undoubtedly have high analytic sensitivity and specificity, it is more appropriate to consider the sensitivity (and to a lesser extent the specificity) of the entire surveillance system. There are a number of sequential steps in the system that may result in false-negative results. The only question is what is a realistic estimate of the average sensitivity of the BSE across all regions of the United States. If the system is only say 50% sensitive, then sample size is underestimated 2-fold. Given the complex testing scheme for verification of BSE, there is minimal likelihood of false-positive results and hence the specificity of the entire surveillance scheme is likely >99.99%.

Evidence from the prevalence analysis (National Surveillance Unit, April 27, 2006) indicates that the prevalence of BSE is less than 1 infected animal in 1 million adult animals with high certainty. Given that this true and was based on a large sample of animals, then this information should be formally incorporated into sample size estimations through a Bayesian approach. There are several published or in-press papers describing Bayesian methods that are relevant to this sample size estimation problem. Johnson et al. (1) considered the estimation problem for a single known infected cluster and Branscum et al. (2) extended this work to include multiple clusters (regions, herds, etc) and allow for zero prevalence in some clusters. These methods should be applicable to the BSE sample size calculation problem with slight modification. The sample size calculations are based on a predictive approach. The Bayesian approach could allow for downweighting of existing prevalence data over time (time value of information). Arguments can be made from a Bayesian perspective that additional testing will not substantially change the belief that the prevalence of BSE is less than 1 infected animal in 1 million adult animals.

In selection of an appropriate sample size, it is important to consider costs (direct costs and opportunity costs) and benefits of testing including those related to trade. Ideally, this should be done using decision analysis. There is no indication that public health benefits in the United States accrue from testing. From a public health viewpoint, the required sample

size is zero. The justification for a non-zero sample size can only be based on perceived trade benefits from testing.

Additional Comments

The second purpose of the ongoing BSE surveillance is purported to be to "provide mechanisms for detection of rising BSE surveillance among U.S. cattle". What criteria will be used to make this determination – point estimates, 95% confidence intervals? The sample size calculation to achieve this goal is not provided. For such a rare event, the sample sizes to actually show this statistically will be huge and cost prohibitive. If this purpose is retained in the document, then the relevant sample size calculation should be included.

References

1. Johnson WO, Su C-L, Gardner IA, Christensen R. Sample size calculations for surveys to substantiate freedom from infectious agents. Biometrics 2004; 60: 165-171.

2. Branscum AJ, Johnson WO, Gardner IA. Sample size calculations for disease freedom and prevalence estimation surveys. Statist Med 2006 (to be published in August).

Review of "Sample Size Estimate for BSE Ongoing Surveillance"

by

Roger S Morris MVSc, PhD, F Amer CE, FACVSc, FRSNZ, CNZM
Co-Director, Massey University EpiCentre, New Zealand

Professor Roger Morris is an Australian veterinary epidemiologist who is Director of the Massey University EpiCentre, in Palmerston North, New Zealand, which has about 70 people involved in teaching, research and consultancy on the epidemiology and control of diseases throughout the world. Current activities include avian influenza in Asia, BSE, foot and mouth disease, development of new approaches to disease surveillance for domestic animals and wildlife, and food safety. Professor Morris has held this position since 1986, and was previously Assistant Chief Veterinary Officer of Australia (1976-1981) and Professor and Chairman of the Department of Clinical and Population Sciences at the University of Minnesota USA (1981-1986).

Professor Morris is a registered veterinary specialist in both epidemiology and pig medicine, with 40 years of experience in these two fields and in cattle disease. He has taken an active interest in BSE since shortly after its initial discovery, and twice acted as an external reviewer of the BSE epidemiology research program of the UK Veterinary Laboratories Agency, Weybridge in the early 1990s. From 1996 to the present, the British government has continuously funded a substantial research program on BSE epidemiology at the EpiCentre, which has examined a range of aspects of the disease. In 2003-4, the European Commission funded the Centre to develop methods for analyzing and interpreting BSE surveillance data. The statistical methods developed are now being applied in Europe, and have formed the basis for current surveillance guidelines adopted by the World Organization for Animal Health (OIE).

Professor Morris was an adviser to the Phillips BSE Enquiry in Britain, and a member of the Food Standards Australia New Zealand (FSANZ) expert group on BSE. He is a FSANZ Fellow, advising the organization in various aspects of food safety risk management. He is also Chairman of SaFoodChain, a multi-country group researching risk-based food safety. He has advised widely on BSE matters.

Introduction

This peer review has considered the proposed sample size and the ongoing surveillance plan, as provided. It considers the issues specified in the revised charge to peer reviewers, in the light of Chapter 2.3.13 of the OIE Terrestrial Animal Health Code entitled Bovine Spongiform Encephalopathy, and associated Appendices 3.8.4 and 3.8.5. The pages of the document are not numbered, so where it is necessary to refer to a specific page, page 1 is taken to be the first page of text, not the cover page.

The task of determining an appropriate sample size for ongoing BSE surveillance presents epidemiological and statistical challenges, because the objective is to make inferences about the state of the "standing population" of live cattle, but sampling can only be conducted on animals which are dead. To achieve a statistically and epidemiologically valid inference from the dead animals back to live ones, requires a special process, which makes use of the fact that in BSE most exposure occurs in calfhood, rather than throughout life, and disease can therefore be considered in relation to birth cohorts of animals, which are born in the same year and share to some degree a similar exposure opportunity. As seen below, this is important in judging the adequacy of the proposed surveillance plan and sample size.

The principles of prevalence estimation for BSE

The accepted calculation procedure uses the spreadsheet program BSurvE to carry out an iterative solution by the method of moments of accumulated surveillance data to estimate either the prevalence of BSE in a country which is known to be infected, or an upper confidence limit on the prevalence which would be consistent with the surveillance data available for a country with very few or no cases. The prevalence within individual birth cohorts is also calculated, with confidence intervals. Because different sub-groups within the population have different value in supporting the estimate, a points system is used to combine the evidence from the different sub-groups in order to assess the adequacy of the surveillance data. Categorization criteria to form sub-groups can include age, surveillance stream, production purpose, geographical area, etc.

The US Department of Agriculture produced a simplified adaptation of the full calculation procedure, which was adopted by the OIE and forms the calculation procedures laid down in Appendix 3.8.4 of the OIE Terrestrial Animal Health Code. It is therefore surprising for the document to suggest that the procedure in the Code is less appropriate for the US than for other countries, and a review of data in the documents which precede this one (Prevalence Estimate, Ongoing Surveillance Plan) do not support the claim that the points requirement is excessive in relation to the cattle population structure and BSE exposure of the US.

Method of calculation used for sample size

The calculation of sample size used in the document does not start from the epidemiological principles which should apply to such a procedure, but rather from the assumption that previous surveillance over the last seven years (March 1999 to March 2006) has provided 6,745,010 points from 735,213 cattle sampled through the four surveillance streams. It is concluded that the average value of these samples is 9.5 points. If samples are collected from the same mix of animals in future, the document concludes that 43,747 samples per year will meet the requirement for reaching 95% confidence that the prevalence of BSE in US cattle is less than 1 per million.

There are a number of defects in this very simple method of estimating the sample size as discussed below.

The numerical calculation: It is stated that the number of points required over 7 years is 2,900,000. In BSurvE, the number of points required is 2,995,730, but OIE rounds the numbers up, so the figure would be 3 million to give 95% confidence that the prevalence is less than one in a million. The number of points is wrongly stated in the document under review to be 2,973,804 and is wrongly attributed to the OIE Code Chapter, instead of to BSurvE. It seems likely that this different number is due to taking a prevalence of 1 in a million to mean 42 infected animals in a an at-risk population of 42 million, ie effectively taking the population at risk as finite and precise. However the population being monitored is dynamic over the seven years it was being sampled, and considerably more than 42 million cattle will have been in the population over that period. BSurvE (and the OIE code) interprets 1 in a million as meaning that each animal has this probability of being infected, and calculates the target based on the Poisson approximation to the binomial distribution. The correct figure is therefore the one in BSurvE.

The document also states (on page 2) that "during the 7 consecutive years prior to March 17, 2006 the US collected 735,213 BSE samples ….and accumulated 2,973,804 points (APHIS 2006b)" but on page 3 states "The prevalence analysis …..reports 6,745,010 points resulting from 735,213 samples (APHIS 2006a)". Such inconsistencies need to be addressed to help the credibility of the document.

If all the assumptions in the document were to be accepted without question, then using BSurvE the number to be tested is 45,049 and using the OIE simplified approach it is 45,113. It is inappropriate to begin the calculation by under-stating the requirement, especially since the number is later further reduced based on weak assumptions.

Random sampling in relation to BSE: The document relates the surveillance undertaken in accordance with the OIE Code Appendix to random sampling. Considerable care is required in making such statements in relation to BSE, and the document does not make its intended meaning sufficiently clear. It is not possible with BSE surveillance to choose animals for sampling from a sampling frame as it is with other diseases, and the concept of "random sampling" from the population at risk is not directly applicable in this context. Random sampling could also be considered within the surveillance streams, but sampling is largely opportunistic rather than random. In any case, for BSE it is better to selectively sample high-risk sub-groups (non-randomly), and make inferences from the sample mix to the population at risk. The claim that "each analytical point calculated by BSurvE corresponds to a single random sample" is not true - a random animal selected for sampling at slaughter will generate 0.37 points (or similar), because this is the estimated probability that an infected animal would leave the herd at a detectable stage – which will be country-dependent as well. Use of the term random sampling in relation to BSE is best avoided, because it does not help understanding.

The difference between detectable and true prevalence: The document makes various statements about detectable prevalence, such as in the last paragraph on page 1 where a statement is made about "the design prevalence of 1 detectable case per 100,000 adult cattle …….". In other places statements appear to relate to true prevalence. Only

about 40% of BSE infected animals are "detectable", because the rest are no longer alive at the age when they would have reached the late stage of the disease at which they would have been detectable by current tests or by clinical examination and pathology.

Both BSurvE and the OIE Appendix are estimating true prevalence, and allow for the fact that not all of the infected animals will be detectable. All reference to "detectable" prevalence should be removed, since it is incorrect in the context of the document.

The age distribution and age at testing of US cattle

RTI Note: Dr. Morris provided additional comments on (1) the importance of more accurate age measurements of the animals and its implications on the prevalence estimation, (2) changes to the required sample size based on the samples collected in preceding seven years; and (3) implications on the sample size if any BSE cases were identified in future years. Although the above discussion may be interesting, it is beyond the charge provided to the reviewers. Therefore, we provide these comments as an Annex to this report in order to improve the readability and organization of the report and to maintain the integrity of the review process. Below we summarize these three points for the benefit of the readers.

- *The worth of each sample in terms of analytic points depends on the accuracy of the age of the tested animals. There can be some issues with the age distribution of animals that is used in previous BSE prevalence estimation. Therefore, a better age data is needed in future.*
- *The OIE requires testing data from preceding seven years to be considered while estimating prevalence. Based on the sample size of data collected in each year, the requirement for the current year can differ. For example, the number of samples required for 2006 may be considerably less than 40,000 samples.*
- *The surveillance plan needs to discuss the implications of identifying future BSE cases because such a find can affect prevalence estimate as well as the needed sample size.*

Sampling strategy proposed to raise the detection power per sample collected

The changes in sampling strategy proposed in the document principally consist of sampling animals which have a higher probability of being detectably infected with the BSE agent than healthy slaughter stock, especially clinical suspects and casualty slaughters. This is a valid objective, and considerable effort should be put into obtaining samples from live cattle with nervous system disorders (i.e., true clinical suspects).

Payment incentives have achieved this in some countries. It is not however sufficient to merely state, as the document does, that this will be achieved. Specific measures will need to be taken to increase the number of samples in these two categories, and to ensure that the age of these animals is known as accurately as possible. Based on UK experience, clinical suspects tend to be younger than the average age so far reported for clinical suspects in the US. The method by which animals from other surveillance streams are

reallocated to clinical suspect should be made more transparent, so that it can be interpreted with confidence.

Further emphasis should be given in the plan to focusing special effort on high-risk sub-populations of animals, selected both geographically and on the basis of potential exposure to contaminated feed material as calves. BSurvE provides a mechanism for incorporating data from such targeted sampling into the calculation procedure in a statistically sound way, but no mention is made in the plan of seeking to use this approach in any suitably structured way.

Given that only two BSE cases have been detected, there is no hard evidence from surveillance so far as to which streams are most likely to contain any BSE cases under US conditions. It is valid to draw samples from surveillance streams in such a way as to maximize points (based on both age and stream), as the plan proposes to do, and this means giving special emphasis to obtaining clinical suspects and casualty slaughters, but also sampling fallen stock where possible, especially in any areas considered higher risk.

Major emphasis in future sampling should be given to obtaining more accurate age estimates on animals, so that points allocated accurately reflect the true surveillance value of each animal. It may be necessary to develop novel methods of ageing animals, such as the use of tooth calcium deposition rings, which are used to age wild animals more accurately than visual methods permit. Alternatively, the adoption of a national animal identification system (as currently proposed) may achieve the required improvement.

As the document states, it is possible to adjust the sampling strategy progressively over the next few years, but this must be done by taking account of accumulated evidence over the seven most recent years, as well as current test results.

The nature of the surveillance objective

The level of sampling required to maintain OIE Type B surveillance in the US is very much lower than that required "to detect BSE at 1 infected animal per million adult cattle in the population with a high degree of confidence" (document, page 1 of text). It is likely that the level and type of surveillance proposed would reliably meet Type B guidelines, and possibly Type A. However these two guidelines are far less demanding than the requirement for showing that the prevalence is less than 1 per million, and if this demanding objective is to be met during ongoing surveillance, the sampling plan must be precisely defined to achieve it. As shown above, the current plan does not have the required level of precision.

Overall response to the charge to Peer Reviewers

RTI Note: We only renumbered the "'overall response" of Dr. Morris so that the first three bullets address the charge questions specifically whereas the fourth point is indirectly relevant to the charge.

1) I expect that the plan as currently proposed would meet OIE Type B surveillance requirements and possibly Type A requirements, but I cannot directly check this with the information I have available.

2) I suggest that the plan should be strengthened by emphasizing the importance of accurate ageing of animals. This could be achieved by various means, including adoption of national animal identification (as currently proposed).

3) I suggest that specific emphasis be given to high risk sub-populations of the cattle population, identified on geographical and risk factor grounds, to strengthen the surveillance plan.

4) The sample size estimate of 40,000 to 43,747 BSE tests per year "to achieve a high degree of confidence that the prevalence is less than 1 per million" is not soundly based on epidemiological and statistical principles and is invalid for reasons described above. It is simply an arithmetic calculation which extrapolates from past testing to future requirements. With the data provided, it is not possible to determine whether this number of samples could be expected to demonstrate that the prevalence of BSE is less than one infected animal per 1 million cattle in the US population, with 95% confidence. For the reasons described above the sampling intensity proposed is almost certainly less than would be required to achieve this surveillance objective. However the OIE requirement is for evaluation over the most recent seven years, and because the US had a peak testing volume in 2003, it is possible that the US might for the next few years meet the requirement with less than the proposed number of tests by virtue of the accumulated sampling over recent years. The data which would allow this evaluation to be made is not publicly available, and it is therefore not possible to judge whether the proposed number of tests would on balance be adequate. In any case, the required number of tests will vary considerably over the next few years, because the accumulated number of points over the most recent seven year period will be heavily influenced by the changing testing intensity of the last few years.

Annex – Additional Comments beyond the Charge to the Reviewers

The age distribution and age at testing of US cattle

If valid conclusions are to be drawn about the BSE status of a national population from inferences drawn using BSurvE or the simplified procedure laid down in Appendix 3.8.4, it is essential that the age of animals tested through each of the surveillance streams be estimated as accurately as possible, especially for high value streams such as clinical suspects. The accuracy of the ageing should be assessed by testing for internal validity, and by comparing with other countries to ensure that there are no obvious defects in the procedure.

The nature of population turnover in a breeding cattle population can only vary within a relatively narrow range because of the nature of demographic processes, especially since the population is derived from only half the animals born, because the proportion of males in the population drops off very rapidly over the first few years of life. Hence it is possible to evaluate whether age data on animals is biologically plausible.

The demographic data which provides the basis for the calculation of a point value of 9.5 per sample is derived from Table 1 of the document "An Estimate of the Prevalence of BSE in the United States" and Table A2 of Appendix A of that document. The internal validity of this data can be checked to determine whether a point value of 9.5 is defensible, since it is highly dependent on whether the ages of animals at testing are accurately estimated.

In the course of producing the estimate of 43,747 animals to be tested, it is stated that turnover in the US beef breeding herd is much slower than in other countries, and hence a higher proportion of animals reach ages where BSE would be detectable by testing. This can also be tested against the evidence.

Table 1 of the Prevalence Estimate document lays out the test results by age group. Since the US does not identify individual animals, ages in Table 1 are estimated. There are a number of significant concerns about the robustness of the allocation of animals to individual cells within Table 1, and misallocation will substantially affect the points allocated, the estimate of prevalence, and the confidence limits on that estimate.

Issues of concern include the following:

1. When people guess ages of animals, especially if they are not very skilled and have only limited background information to assist them, they are likely to guess values in the middle of the likely range, and over a large population this will show up in the form of over-representation of some age groups and under-representation of others. In this table, the ages 5 and 10 years both have marked peaks within the distribution compared with adjoining ages, and 15 has a small peak. There are also no animals of unknown age in the data set.

The distribution by age is remarkably similar between the different surveillance streams, whereas in other countries each stream shows a different age distribution. This is shown in

Figure 1. It is of note that the peak of fallen stock seen in 2 year olds in the UK is missing in this data, and fallen stock are exceptionally high in the middle age ranges, compared with other countries. As can be seen, fallen stock in the age range 4 to 6 years make up a very high proportion of total tests, and each of these animals contributes 0.9 points to the total, compared with 0.2 for animals under 4 years and 0.4 for animals 7 to 9 years.

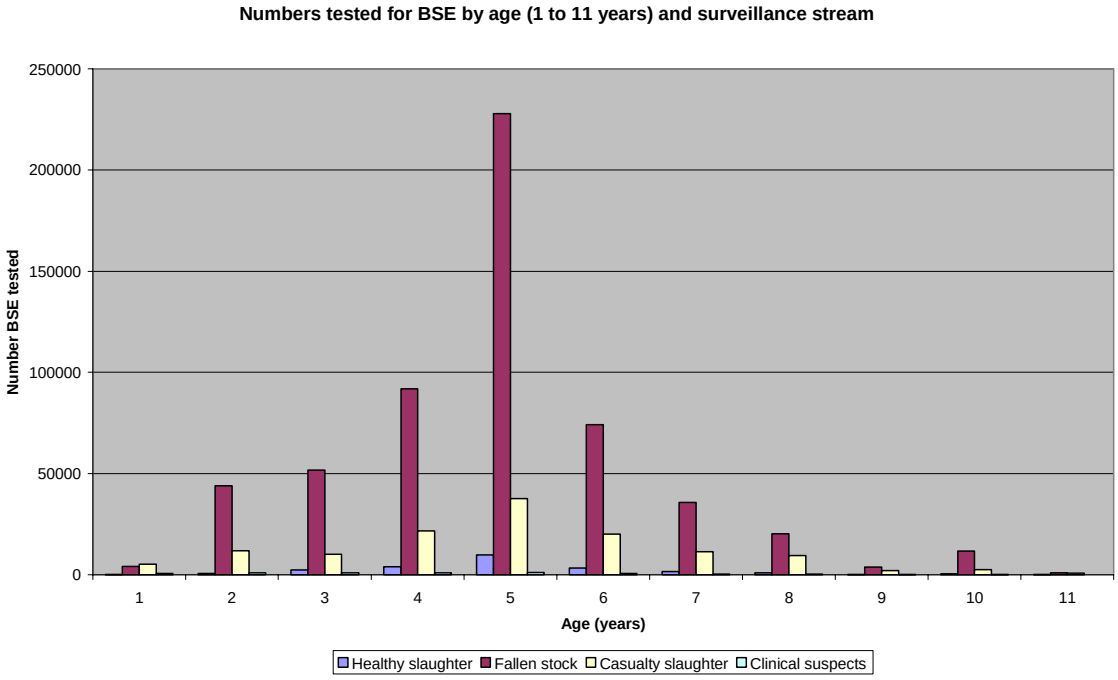

Figure 1. Distribution of BSE tests 1999-2006 by age group (1 to 11 years) and surveillance stream

Table A2 provides data on which to calculate the percentage of animals removed in each different age group, and this can be compared with the animals tested for BSE in each age group.

If animals less than two years old are excluded from the comparison, the difference is still very apparent, as shown in Figure 3.

Even allowing for whatever efforts were made to select from the age group 4 to 6 years, the discrepancy between the proportion of animals available to test in the age group and the proportion of tests occurring in the age group is extreme, and the over-representation of five year olds seems unlikely to be true.

2. Animals in the 4 to 6 year age group at testing gain the highest point score and contribute most to the prevalence assessment. If animals are misallocated to this age range, they will inflate the points value very substantially, and therefore influence the estimation procedure. Due in part to the peak at 5 years, the age range 4 to 6 years

comprises 73% of healthy slaughters tested, 69% of fallen stock, 59% of casualty slaughters, and 42% of clinical suspects. These are extremely high percentages for three of

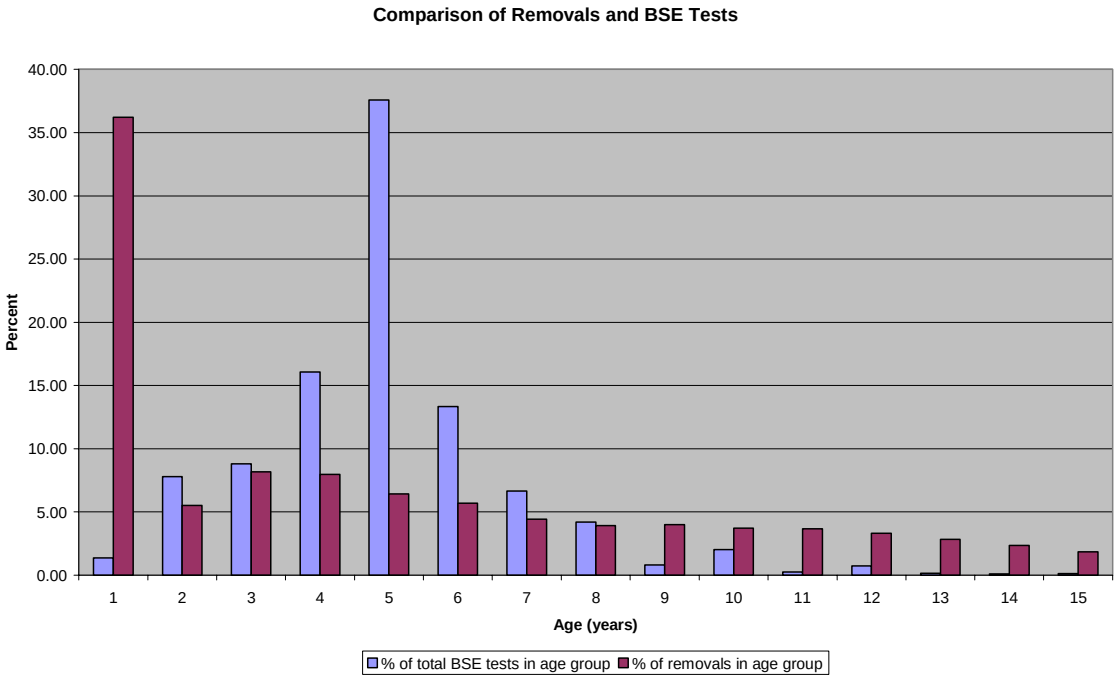

Figure 2 Percent of animals removed from each age group and percent of all BSE tests in the age group

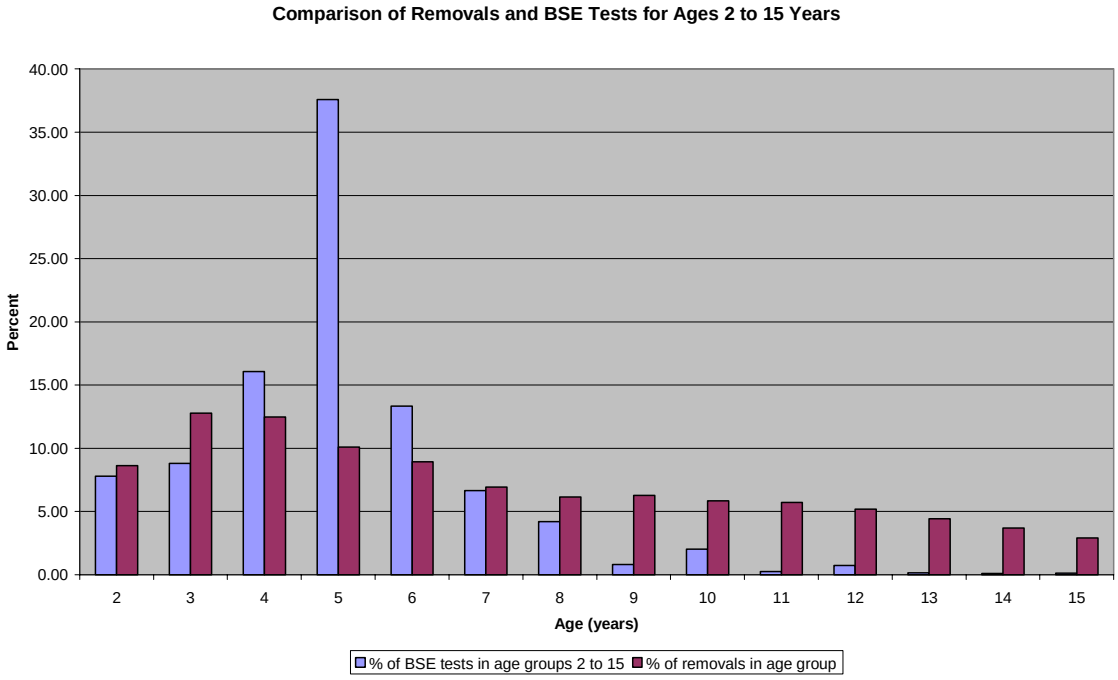

Figure 3 Percent of animals removed in the age range 2 to 15 years, compared with BSE tests in those age groups

18 age values. The most recent UK data (2003) gives 75% of healthy slaughters in the same age range (suggesting a similar age of normal culling in the two countries), but 28% of fallen stock in the age range, 34% of casualty slaughters, and 27% of clinical suspects. Since the last three groups have high point values compared with healthy slaughters, substantial misallocation on age to this range will seriously over-inflate the points score.

3. Some groups are also under-represented – for example fallen stock in the 2 year old age group (typically a very high risk group to die on-farm) and also in the oldest age group, where again the risk becomes higher. It is typically low in the middle years.

4. The shapes of the frequency distributions for the four surveillance streams are not identical, but are far more similar than would normally be expected, suggesting that the age estimation procedures used were seriously flawed.

5. The document being reviewed makes the unsupported statement that beef cows in the US will typically stay in the herd until much older ages than is the case in other countries, and therefore be available to express clinical BSE for a longer period than cattle in other countries. However mean ages at death can be calculated from Table 1 of the Prevalence Estimate and are compared with data for the UK in 2003 (in brackets). Healthy slaughters are culled at a mean age of 5.2 (4.9), fallen stock at 5.0 (6.0), casualty slaughter at 5.1 (6.4) and clinical suspects at 4.3 (2.8). Ages at death in the various US surveillance streams are much more similar to each other than would be expected, and do not show the same variability as do the equivalent UK streams. The US data on herd life in Table A2 is based on extrapolation from very limited data, especially for beef cattle. Calculation of mean age of animals in the national herd from the Table gives 5.2 years for beef cattle, 2.1 years for dairy cattle, and 4.3 years for the total population. This is shorter than for various European countries examined (Portugal is particularly long) and does not provide greater opportunity for expression of BSE signs or for animals to be positive to the various tests.

6. Because few clinical suspects for BSE were reported in the US, animals from other streams were reallocated to this stream retrospectively if their reported clinical signs were consistent with BSE – in other countries these animals would have remained in their original streams. While this is defensible if the evidence supports it, the use of it in the Prevalence Estimate document is so non-specific about the criteria used for reclassification that it is not possible to judge whether the procedure was epidemiologically valid. It is important that the exit probabilities (which determine the point values) reflect the stream definitions.

It is concluded for this assessment that the allocation of animals to age groups up to 2006 has almost certainly raised the point score for past testing substantially above its true value. In order to provide continuing assurance that the prevalence is below 1 in a million, it is essential that better age data be obtained in future, and the estimate that each animal tested will on average be worth 9.5 points is over-optimistic for selecting a sample size.

Period of time to which the prevalence estimate applies

The implication of the sample size estimation procedure used is that all that is required is to calculate a mean number of animals to be tested per year. However the OIE requirement is for data to be considered over a seven year period. Each year the oldest data is dropped off, and the newest year added. The use of a fixed annual testing volume is only valid if past testing has been at a constant rate.

The USDA web site provides data up to 2004 on the testing history, as shown in Figure 4. This shows very clearly that testing has not been constant, but rose sharply from 1999 to peak in 2003. Therefore when 2006 testing replaces 1999 testing in the calculations, the number of samples required for 2006 may in fact be considerably less than 40,000 – this requires a proper calculation in accordance with the OIE Appendix to determine, and requires access to the annual testing data split by age group and surveillance stream. This is not provided on the USDA web site.

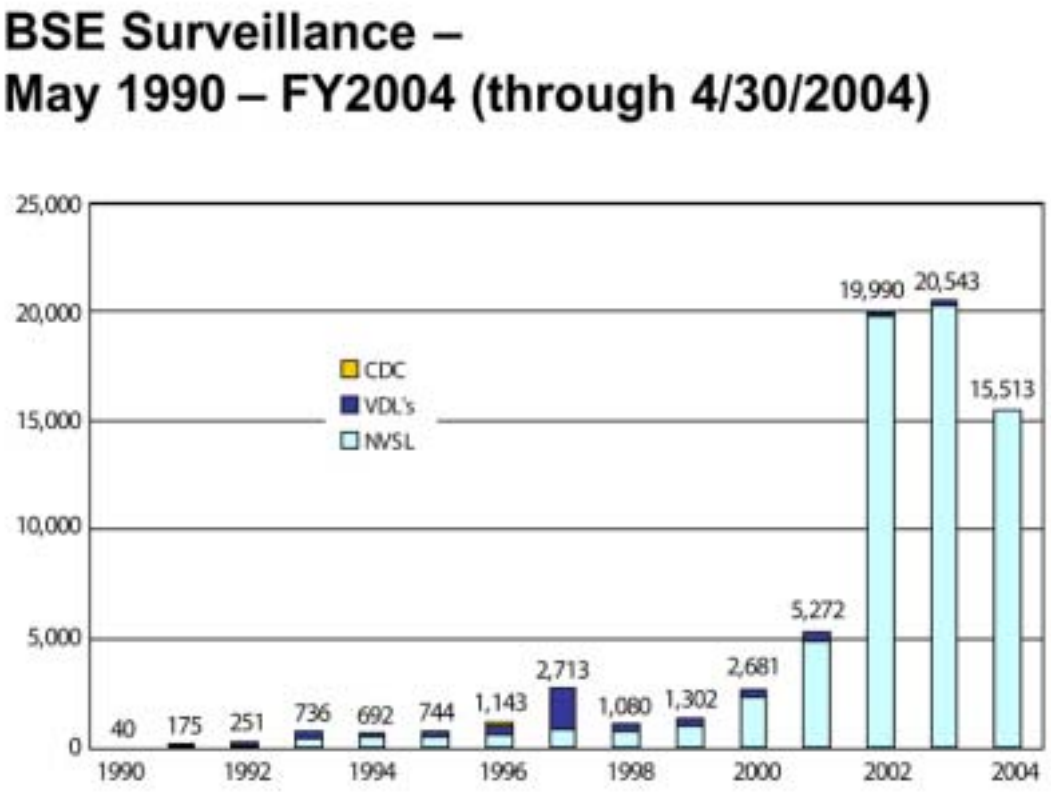

Figure 4 BSE testing data by year 1990-2004

The implications of finding any further cases

So far two autochthonous BSE cases have been found in the US. Both animals were of uncertain age, estimated to have been approximately 12 years old (Texas case, 2005) and approximately ten years old (Alabama case, 2006), according to the two final epidemiology reports. Incubation periods of this length occur in the UK, but most incubation periods are

considerably shorter. To find two cases with incubation periods of this length implies that there was almost certainly a larger number of earlier cases with shorter incubation periods which were not detected, and from which infectious material may possibly have entered the feed chain. Cases arising from any recycling of infection in this way would be expected to be detectable after a further incubation period had elapsed. In designing ongoing surveillance, this possibility must be given consideration.

This age estimation places the two known positives in cohorts beyond the seven years covered by OIE surveillance requirements, and therefore the US is considered to have had zero cases within the last seven birth cohorts. Should any further cases be detected by ongoing surveillance, the amount of testing required to support a claim of prevalence below 1 in 1 million would rise very sharply, with the requirement depending on how many cases are found.

However it should be noted that if these two animals (estimated birth cohorts 1993 and 1996) are to be excluded from consideration, all negative animals born before 1997 must also be excluded from the points total, because of the nature of the epidemiological calculation based on birth cohorts.

The ongoing surveillance program as planned relies on all future test results being negative. It would be wise to give consideration in advance to the implications for the surveillance objective and sampling strategy, should any positives be detected.