Author's response to reviews

Title: Diagnostic accuracy of cerebrospinal fluid protein markers for sporadic Creutzfeldt-Jakob disease in Canada: a 6-year prospective study

Authors:

Michael B Coulthart (michael.coulthart@phac-aspc.gc.ca)
Gerard H Jansen (gjansen@ottawahospital.on.ca)
Elina Olsen (elina.olsen@phac-aspc.gc.ca)
Debra L Godal (debra.godal@phac-aspc.gc.ca)
Tim Connolly (tim.connolly@phac-aspc.gc.ca)
Bernard CK Choi (bernard.choi@phac-aspc.gc.ca)
Zheng Wang (zheng.wang@phac-apsc.gc.ca)
Neil R Cashman (neil.cashman@vch.ca)

Version: 2 Date: 21 September 2011

Author's response to reviews: see over

September 21, 2011

Dear Mr. Bongcayao,

We are grateful for the Journal’s prompt and thorough review of our manuscript, and for the kind comments and insightful criticisms provided by the reviewers. We also welcome a review of our statistical analyses, and look forward to receiving any additional comments.

A revised version of the manuscript has now been submitted to the Journal, in MS Word format. All changes in the text, tables and figure captions are indicated with Track Changes. Links between reviewers’ comments and specific passages of text have also been highlighted and labeled, using MS Word Comments. To facilitate review of our proposed responses and revisions, we have reproduced the reviewers’ original comments in their entirety below with our responses (including descriptions of revisions) and supporting discussion placed immediately below each specific comment in a contrasting font. Because some of Reviewer #2’s comments contained multiple points but were not numbered separately, we have in these cases separated and numbered the individual points in the series i, ii, …. Where page and line numbers are cited by us below, they refer to the revised version of the MS; we trust that this notation is easy to follow.

We would also like to note that even after careful consideration we found ourselves in disagreement with some of the reviewers’ specific comments, and could not clearly see an opportunity to respond with a specific revision of the manuscript. Each of these cases is clearly noted, and accompanied by supporting discussion explaining our point of view. If the Editor and/or reviewer disagree with any of our proposed responses, we would of course be more than happy to clarify further or to reconsider.

A. Editorial comments

Journal requirements regarding the non-necessity of keywords, as well as counts of figures, tables and words, have all been addressed (see Title page). In addition, all internal vertical lines have been removed from tables (see pages 37-39).

B. Reviewer #1 (M. Tranulis)

Reviewer’s report
Diagnostic accuracy of cerebrospinal fluid protein markers for sporadic Creutzfeldt-Jakob disease in Canada: a 6-year prospective study

This contribution by Coulthart and co-workers examines the diagnostic usefulness of spinal fluid proteins (14-3-3, tau & S-100B) for sporadic CJD. This is an area of considerable interest and with some variations between reports from different labs.

This is a well-composed and well-presented contribution that will be of interest in this field. The authors particularly advocate the use of likelihood ratios (LRs), which is elaborated on in the discussion, which is almost exclusively dedicated to statistical analysis/discussion.

This is certainly important and of great interest for the reader, therefore, the editors might find it useful to have a separate expert review of this part of the study. To this reader, which is not a bio-mathematical expert, the statistics and presentation and discussion of the data seem correct, balanced and reasonable. Figures and tables are easy to interpret and the use of confidence intervals is highly welcome.

Response: We kindly thank the reviewer for his supportive comments.

Comment 1: The authors are advised to use the term non-CJD (as they do in the abstract and fig 1) and not nCJD, which might cause unnecessary confusion, being similar to sCJD, vCJD, gCJD etc., which are abbreviations for sub-groups of CJD.

Response: We agree. All instances of the term “nCJD” in the MS have been replaced with the term “non-CJD”.

Comment 2: It is well documented that the proteins in question, 14-3-3, tau and S-100B are pretty robust and stable, allowing reliable measurements even after prolonged sub-optimal storage. Nevertheless, this reader finds the materials and methods part dealing with sample storage (temperature, duration etc.), preparation/extraction, any F/T cycles, master standards for calibration (fresh?, storage?, how and for long?, stored diluted?), to be an area the authors might have a look at. Correct treatment of samples and reagents is of pivotal importance to insure data integrity – errs in this department cannot be corrected by statistical analysis. For example, in this study 14-3-3 is analyzed first and then, based upon sample availability and the outcome of the 14-3-3 analysis, S-100B and tau are analyzed – how was the sample stored during this time? Frozen or at 4°C? Some readers, including this one, would find this information to be interesting and something that would contribute to the completeness and quality of this paper.

Response: We agree, and have added text (Methods p 10 lines 7-8 and p 11 lines 3-5) to
further highlight the ways in which some of these important technical details were dealt with during the study. We have also added text in the Discussion (p 24 line 18 – p 25 line 7) to support the reviewer’s point that 14-3-3 and tau proteins show good stability during pre-analytic handling (although published studies of this nature appear to be lacking for S100B).

**Discussion:** It is of course critical to optimize and standardize as far as possible the pre-analytic handling of samples and reagents used in protein immunoassays, to try to ensure that analytic results are not compromised by loss of integrity of labile analytes and chemicals. A dedicated research laboratory working within a well-defined study protocol and perhaps within a single center can exert significant control over such factors. However, a reference laboratory can advise submitting laboratories on optimal procedures but it cannot control or even fully define the conditions to which specimens are exposed during pre-submission collection, handling, storage and shipping.

On the other hand, we think that the excellent diagnostic performance we observed for the three CSF markers we studied despite this incomplete control of pre-analytic variation is a strength of our study – in other words, if the markers are diagnostically robust under sub-optimal conditions, their performance would only be expected to be better if conditions were more fully controlled. As the reviewer notes, both 14-3-3 and tau proteins in CSF are also quite stable under laboratory conditions with regard to their reactivities in the particular immunoassays we have used. We have added text to the Discussion and cited some published studies to specifically support this point (Beaudry P *et al* 1999, *Dementia and Geriatric Cognitive Disorders* 10:40; and Schoonenboom N *et al* 2005, *Clinical Chemistry* 51:189). We were however unable to locate similar published studies for S100B.

A considerable length of text would be required to completely specify our own post-receipt laboratory procedures in the MS, and in the original version we approached this need in part by referring to the excellent manuals provided by the ELISA kit manufacturers and by noting any key points that are not specified therein (Methods p 9 line 23 – p 10 line 2). We have however enhanced the description of post-receipt handling of CSF specimens, with reference to shipping, freeze-thaw treatments after receipt, and interim storage.

**Comment 3:** References 3 and 26 are incomplete.

**Response:** These formatting errors have been corrected.
C. Reviewer #2 (P. Gambetti)

Reviewer's report

Title: Diagnostic accuracy of cerebrospinal fluid protein markers for sporadic Creutzfeldt-Jakob disease in Canada: a 6-year prospective study

Version: 1 Date: 17 August 2011

Reviewer: P. Gambetti

This is a well performed study aimed at characterizing the performance of three CSF markers, 14-3-3, tau, and S100B. The narrative suffers somewhat by deviating from the central theme, some repetition, and by omitting important simple detail. The paper relies heavily on statistical modeling and shows a preference for likelihood ratios. This satisfies a more statistically inclined reader. However, likelihood ratios have failed to gain wide acceptance in general clinical practice despite frequent recommendations. Nevertheless, with appropriate modification, the manuscript provides valid and useful information.

Response: We kindly thank the reviewer for his supportive comments. We agree (along with many other authors, some of whose work is cited in the manuscript) that the use of likelihood ratios has not yet penetrated sufficiently into clinical practice, and the reasons for this constitute an interesting subject in itself. We hope that our contribution helps in some small way to close this widely acknowledged gap between theory and practice, at least in the sub-speciality area of CJD diagnostics.

Comment 1: P9 ln17 # P10 ln 4. Standard associated procedure with too much detail needs to be shortened.

Response: We feel this level of procedural detail is essential to the MS. Although further discussion is welcome, we would prefer to leave the original text unchanged.

Discussion: These two short paragraphs (total 117 words) contain essential technical detail regarding how we performed our neuropathology examinations and molecular genetic analyses, including descriptions of some reagents that are often selected or formulated differently by various expert CJD diagnostic laboratories. The 12F10 anti-PrP antibody for example is not used by all laboratories performing CJD immunohistochemistry. Also, the two oligonucleotide primers used for PCR were designed de novo by the Canadian CJD Surveillance System (see Acknowledgements, p 28 lines 8-9) and are, to the authors’ knowledge, both unique to this laboratory and previously unpublished. For these reasons we feel the detail should be retained, so that readers may fully understand (or even emulate) our technical approach, and others could in principle reproduce our findings.

Comment 2: P10 ln20.
i) Were tau and S100B assayed with the same frequency as 14-3-3?

Response: We believe this point was addressed in the original MS. Although further discussion is welcome, we would prefer to leave the original text unchanged.

Discussion: As noted in the original MS (Methods p 10, lines 21-22 and p 11 line 18 – p 12 line 7), the ELISA assays were performed on all samples for which this was technically feasible and specific patient-exclusion criteria did not apply. The numbers of samples among the 1000 tested for 14-3-3 that could ultimately also be tested individually by the two ELISA assays (926 and 924 for tau and S100B respectively), as well as the number that were tested jointly by both ELISA assays (913), were also given in the original MS (Methods p 12, lines 7-8; Results p 14 lines 22-23).

ii) Were controls included in these quantitative assays? If yes, what were the precision statistics, mean, CV%, number of values.

Response: We have clarified these important technical details by adding new text (Methods p 11 lines 5-11).

Discussion: The general problem of identification, formulation, characterization and application of reference materials that are suited for use in quality control of quantitative immunoassays of protein analytes is a complex one. In the case of CSF proteins, this problem has not yet been solved even for very well-studied markers such as tau and Aβ as they are applied to diagnosis of Alzheimer’s disease (see for example Mattson N et al 2011, Alzheimer’s & Dementia 7:386). Some laboratories take pragmatic approach to obtaining control materials by simply selecting surplus quantities of appropriate samples from their ongoing work stream. However, a significant practical problem with this approach is that specimens with informative levels of reactivity – i.e., those falling just above or just below assay cutoff thresholds, as unreactive or highly reactive specimens do not provide much quantitative information on stability of a method – occur only infrequently. Even when such informative samples are identified, the small CSF volumes provided by submitting laboratories are usually not sufficient to support their long-term ad hoc use as internal controls. These limitations apply to any CJD testing laboratory, and even more so for one such as ours that receives only a modest number of samples. The need can be addressed to some extent if control materials are provided with the ELISA kit; although this is the case for the S100B kit we used, for the tau kit unfortunately it is not. We are currently discussing with the manufacturer of the tau kit (Innogenetics) whether and how this technical gap might be filled.

Comment 3: P11 ln9 #19 and P13 ln3 #12. These narratives and Figure 1 provide repetitious information.

Response: We do not agree that the cited text and Figure 1 are redundant. Although further discussion is welcome, we would prefer to leave the original text and figure unchanged.

Discussion: Our presentations of inclusion and exclusion criteria (Methods p 11 line 18
– p 12 line 7) and the actual numbers of patients who were classified on the basis of these criteria (as per Figure 1) are not conceptually the same, and we feel that to fully and clearly describe our approach to case selection we require both. An option for revision that we considered was to include the cited text in the caption of Figure 1, but this does not significantly reduce overall bulk of the MS, and in our opinion the relevant text still seems most appropriately placed in the Methods section.

Comment 4: P12 In2 #8. One short sentence should suffice.

Response: We feel retention of this short passage containing essential technical detail is important for the sake of completeness. Although further discussion is welcome, we would prefer to leave the original text unchanged.

Discussion: These 3 sentences (total 100 words, Methods p 12 lines 15-21) define a widely endorsed criterion of sample size that should be met (as does our study) to help ensure that logistic regression analyses are statistically valid. We also attempted to condense the already brief conceptual exposition, data description and conclusion (sentences 1, 2 and 3 respectively) into a single short sentence, but were unable to do so without sacrificing readability. We therefore feel it is important to retain the explicit detail here, particularly because some readers of BMC Neurology may not be completely comfortable with such nuances of statistical method, and it should not be necessary for them to consult the literature to understand this simple point.

Comments 5a–5d: This section as well as Fig.1 are confusing, raising a number of questions to which the authors should answer.

5a. Why the 29 cases with the diagnosis of “probable” sCJD were excluded in view of the high probability that many of these cases were actually sCJD.

Response: We feel this point was fully explained in the original MS. Although further discussion is welcome, we would prefer to leave the original text unchanged.

Discussion: The reason for excluding the 29 cases with a final diagnosis of probable sCJD from the main statistical analysis of diagnostic accuracy was explained in the original MS (Methods, p 12 lines 4-6). Stated briefly, because a positive 14-3-3 assay result is one of the criteria used to classify cases as probable sCJD, to use cases classified in this manner to assess the diagnostic accuracy of the 14-3-3 test would constitute circular reasoning. This would not be true of course for cases that had been classified as probable sCJD independently on grounds other than a positive 14-3-3 result (e.g., presence of ca. 1-Hz periodic sharp-wave complexes on EEG). However, none of the 29 probable sCJD cases in question were classified on the basis of this independent criterion, and we therefore felt we had no option but to exclude them from the main analysis.

5b. Is it fair to exclude “probable” sCJD and to keep probable nCJD?

Response: For practical reasons, our study was indeed based on the use of somewhat different diagnostic standards for sCJD and non-CJD cases – autopsy in the former case;
heterogeneous criteria usually not including autopsy in the latter. We do not consider this to represent a logical flaw or a serious empirical weakness of the study. Nevertheless, we have clarified our views on the potential effects of residual diagnostic uncertainty in the “probable non-sCJD” group, by enhancing our presentation of the significance of Canadian CJD surveillance data in this context and relocating it from Results to the Discussion [see Results p 14 line 19 (deletion); Discussion p 24 lines 9-16 (insertion)].

**Discussion:** It is true that our study design is based on comparison of a group of autopsy-confirmed cases of sCJD with a control group in which approximately 95% of non-CJD diagnoses were not (to our knowledge) verified by autopsy. However, we do not see a logical necessity for both study groups to be selected on the basis of equivalent diagnostic standards, and in light of other supporting data we do not believe that this approach constitutes a serious empirical weakness in the study in the form of diagnostic inaccuracy. More specifically on the latter point, in the original MS (Results p 14 line 19, now deleted and relocated to Discussion p 24 lines 9-16) we presented arguments from our Canadian CJD surveillance data that the number of unrecognized cases of prion disease in the non-CJD group is likely to be small (albeit ultimately unquantifiable). Furthermore, because the size of the non-CJD group (873 cases over the 6-year study period) represents a large denominator, we expect the proportional effects of any residual diagnostic uncertainty on our estimates of diagnostic accuracy (from CJD cases being misclassified as non-CJD) to also be small. With this said however, because of the high importance of clarity on this point, this section of text has now been expanded and relocated to the Discussion.

**5c.** If the PRNP was determined in only 72 cases how could it be “definitely” confirmed that 127 cases were sCJD and 5 fCJD? Alternatively, were they definitely prion disease and “probably” sCJD?

**Response:** We acknowledge the validity and importance of this point, and thank the reviewer for pointing it out. We have added text (Discussion p 22 line 20 – p 24 line 7) to (i) clarify the fact that we cannot completely rule out the possibility of a small fraction of genetic prion disease cases among the 55 cases that we classified as sCJD but for which no genetic information was available, and (ii) articulate why we do not expect the residual diagnostic uncertainty to have significant effects on our main conclusions. We have also revised Figure 1 to indicate the lack of genetic information in the above-mentioned 55 cases. Please note also that we corrected some numerical errors in the numbers of cases in the categories under “No neuropathology” located at the lower right of the diagram. These numerical changes do not affect our main conclusions, and required no corresponding revisions of the text. Finally, we added a clause in Methods (p 9 lines 11-12) to more clearly indicate that in addition to administering standard intake questionnaires to give referring clinicians an opportunity to provide contextual information on patients, including family history, the CJDSS also systematically collects information on family history via a direct interview with family members that also includes a standardized questionnaire.

**Discussion:** Please note first that the 5 definite cases of genetically cause prion disease indicated in Figure 1 were all confirmed as such by PRNP sequencing and
neuropathology, and as such were not included in the 127 autopsy-confirmed cases of prion disease used in our study of CSF marker performance (see Methods p 12 lines 3-4). In 55 of the 127 autopsy-confirmed cases however, despite the presence of a CJD-like clinicopathological phenotype and no family history of a similar disease (ascertained through consultation with clinicians as well as family interview) we were not granted consent to perform DNA sequencing of the PRNP gene to confirm the absence of mutations. The presence of PRNP mutations in some of these putative sCJD cases would indeed imply that such cases had been misclassified – this is an example of the problem, widely encountered in empirical studies of diagnostic test accuracy, of an imperfect diagnostic standard. However, the following considerations lead us to believe that the impact of this diagnostic uncertainty (which we note in this patient group would be through effects on a marker’s diagnostic sensitivity) would be small, and therefore would not significantly influence our main conclusions:

(i) Misclassified cases of genetically caused CJD (gCJD) likely constitute a minor component of our designated sCJD group. In general CJD surveillance populations, ca. 10% of human prion disease cases presenting clinicopathologically as CJD (as distinct from the well-known Gerstmann-Sträussler-Scheinker phenotype, or other less frequent presentations, which were excluded from our study) have been estimated to be genetic in origin, being linked to point mutations or DNA insertions in the PRNP gene (Kovacs G et al. 2005, Human Genetics 118:166). Furthermore, a significant subset of confirmed gCJD cases lack a positive family history – for example, ca. 50% of 114 such cases compiled by Kovacs et al. that were linked to the most common pathogenic PRNP allele (E200K) lacked such a history. However, when applied to our subpopulation of 55 cases without genetic information, these considerations lead us to expect only a small number of unrecognized gCJD cases – i.e., fewer than 5 (~10% of 55), which is <4% (5/127) of the total number of autopsy-confirmed prion disease cases). We feel that this is still an excellent level of accuracy for a diagnostic standard, and that the small residual diagnostic uncertainty attributable to misclassification of gCJD as sCJD should not be large enough to detract significantly from our main conclusions. We have therefore provided a discussion of this aspect of our data in the MS, including estimated numbers of unrecognized gCJD cases in our sCJD group, along with confidence limits.

(ii) The quantitative effects of any inadvertently included gCJD cases on statistical estimates of assay performance are expected to be minimal. A recent study of the diagnostic accuracies of these three CSF markers in a series of 174 patients with various genetically caused prion diseases, including 117 with gCJD (Ladogana A et al 2009, Journal of Neurology 256:1620), found that their diagnostic sensitivities in gCJD were comparable to those observed for sCJD both generally and in our study – 0.83, 0.86 and 0.87, respectively. This indicates that even if our designated sCJD subpopulation included a significant number of gCJD cases, the distorting effects on our statistical estimates of diagnostic sensitivity would be expected to be minimal. We have also included a discussion of this aspect of our data in the MS.

5d. Is there an explanation as to why no “possible” cases of sCJD were observed.

Response: This too is a valid point, and merits a small clarification in the MS (Results p
to make it explicit that we did not assign any cases a final classification of possible CJD.

**Discussion:** The formal surveillance case criteria for possible sCJD, which are useful when only clinical information is available, are as follows:

i) progressive dementia of less than 2 years’ duration

and

ii) at least 2 of:
   - myoclonus
   - visual or cerebellar signs
   - pyramidal/extrapyramidal signs
   - akinetic mutism

These criteria differ from those for probable sCJD essentially by virtue of the latter including a positive 14-3-3 assay result and/or presence of a typical EEG displaying triphasic periodic sharp-wave complexes at ca. 1 Hz. Thus, in a population of patients such as ours defined *a priori* as having all been tested for 14-3-3, to qualify as possible sCJD any individual must have (i) met the clinical criteria above; (ii) had a negative 14-3-3 test result; (iii) not displayed a typical EEG pattern; (iv) not undergone neuropathological examination; and (v) not received a more compelling alternate diagnosis.

The Canadian CJD Surveillance System only infrequently assigns a final case classification of possible sCJD on the basis of the above-mentioned criteria, largely because our autopsy rates are relatively high (*e.g.*, ~84% of cases classified as definite or probable prion disease are definite) and because there is careful follow-up with referring clinicians to ensure that information on clinical picture and results of supporting investigations is as complete as possible. In the finite sample of patients on which the present study was based, it is therefore not unexpected that we simply did not encounter any instances in which all of the criteria for possible sCJD were met. It is also important to recognize that while a particular patient may meet the criteria for possible sCJD, the total evidence may simultaneously be more compelling in favor of a specific non-CJD diagnosis, or at least against CJD as the correct diagnosis.

**Comment 6:** P14 In2. ROC curves are an important determinant for comparing assays with both shape and AUC being helpful. These curves should be shown in a single figure.

**Response:** We agree. A new Figure 2 and caption have been added including the requested ROC curves and supporting information, as well as supporting text (Results p 14 lines 22-23). Several subsequent original figures have therefore been re-numbered: original Figure 2 is now Figure 3; and original Figure 3 is now Figure 4. (Note that the original Figures 4 and 5 have been deleted in the revised MS; please also see our responses to Comments 9 and 10 below.)
Comment 7: P15 ln20.
i) This equation has been derived from the data used for the manuscript, and cannot be applied with confidence to other data sets.

Response: Our application of logistic regression was always intended to be restricted to its illustration of a specific aspect of the current dataset. However, we have added a small emphasized phrase in the text (Results, p 16 line 15) to further clarify this.

Discussion: We certainly agree that the logistic regression equation we derived to help assess the joint influence of tau and S100B assay results on the post-test probability of a diagnosis of sCJD is, strictly speaking, limited in its application to the patient population we studied, and cannot be extrapolated uncritically to other such patient populations. However, we do not believe we suggested otherwise in the original MS, or indeed that risk scores derived from the equation could be used directly in clinical diagnosis at all. Instead, our intent was to use the bivariate risk scores derived from the equation (and a related trivariate risk score calculated from a separate logistic regression analysis that included the 14-3-3 assay result as a discrete variable), to compare the overall diagnostic accuracies of the different CSF markers and combinations thereof using ROC-based criteria. The approach and the results of these comparisons were detailed in the original MS (Results, p 16 line 14 – p 17 line 5, and p 17 lines 18-23. Nevertheless, given the importance of the issue and the potential for misinterpretation, we have slightly modified the text to further clarify the original motivations of our use of logistic regression.

ii) That being the situation, the authors need to provide alternative models for other potential users wishing to use the joint markers. Possibilities include: if either tau or S100B is positive, then the case is deemed positive, or both tau and S100B have to be positive for the case to be deemed positive, etc. Each model will have to be tested to find the most appropriate.

Response: Given our discussion under Comment 7i, and that Comment 7ii seems to be predicated on Comment 7i, we feel that the substance of Comment 7ii was already fully addressed in the context in which it was presented here. However, we also felt that the independence of the two approaches we used to assess the effect of combining markers at different thresholds perhaps needed to be better emphasized, and we have added a small amount of text (Results p 17 lines 7-8 and p 18 lines 1-2) to accomplish this. Please note also that the issue raised in Comment 7ii appears to be closely related to that raised in Comment 8, which we discuss further below.

Discussion: It is important to distinguish clearly between the two approaches we used to assess the benefits of combining CSF protein markers and how these benefits vary with choice of cutoff threshold. The first approach, discussed under Comment 7i, was that of logistic regression, which is a technique designed to fit a logistic function to the relationship between a risk score and one or more independent variables, which can be either continuous or discrete. This approach does not require any pre-selection of cutoff thresholds for continuous independent variables, which in our case are the assay values for tau and S100B. The second approach, quite different conceptually from the first, was based on construction
of 2x2 contingency tables using different values of cutoff thresholds for tau and S100B to illustrate their diagnostic power in a number of pre-chosen scenarios. In this light we felt that the reviewer’s Comment 7ii could be addressed most directly by simply clarifying the independence of these two approaches. Please note however that the issue raised in Comment 7ii appears to be closely related to that raised in Comment 8, and we therefore discuss it further below.

**Comment 8:** P16 ln10 #18. This appears to be considering selective data where both tau and S100B are either both above or below the chosen thresholds. How many cases are there where the values are discordant?

**Response:** We do not agree that this constitutes arbitrary selection of data. We welcome further discussion, but at this point we prefer to leave the original text unchanged.

**Discussion:** In the cited section of the MS, we constructed bivariate 2x2 contingency tables and calculated test-performance statistics – particularly likelihood ratios – to **illustrate** the effects of marker combinations and test-result cutoff thresholds on diagnostic power. We therefore chose to present the results for only 1 of the 3 possible joint test outcomes (i.e., both markers above a particular threshold) that could in principle be used to define a “positive” result for 2 markers, as well as for 1 of the 3 possibilities (i.e., both markers below a particular threshold) that could be used to define a “negative” result. We do not feel this constitutes an arbitrary selection of data, but again, an **illustration** of the above-mentioned effects, by choosing two particularly clear examples of how diagnostic probabilities can be strongly modified by these protein markers in subsets of patients with particular assay results. A more complete exploration of the manner in which statistics such as likelihood ratios or logistic regression-derived risk scores vary with marker combinations or threshold values could indeed be undertaken, but we felt that such a study would fall beyond the scope of the present contribution.

Please note also that Figure 2 includes all of our study’s data points (plotted separately for sCJD and non-CJD groups), as well as the locations of several illustrative diagnostic threshold values used for tau and S100B as explained in the text. The reader interested in exploring patterns of diagnostic classification based on alternative test result levels and combinations thereof can obtain a general sense of this by visual inspection of the figure.

**Comment 9:**

i) It is generally accepted that as more of the clinical symptoms of ataxia, dementia, and myoclonus occur there is an increased probability of CJD being present in a patient. Figure 4, if kept, should be placed in the supplement.

**Response:** We agree that these components contribute little new information. Figure 4, its caption (p 42 line 8, deleted) and associated text [Results p 19 line 7 (deleted) and Discussion p 25 lines 15-21 (deleted)] have accordingly been removed.

ii) More useful, and more helpful than figure 2, would be a pair of Venn diagrams showing the three assays as positives and negative cases, with optimal decision points being used.
Response: We feel that such a revision would entail a significant loss of information. We welcome further discussion, but at this point we prefer to leave the original text and figure unchanged.

Discussion: Please note that Figure 2 graphically displays the quantitative distribution of all of our study data, as well as the locations of several illustrative diagnostic threshold values used for tau and S100B as explained and applied in the text. One of the key conclusions supported by our data is that systematically taking into account the quantitative results of CSF protein marker assays can offer the clinician uniquely valuable diagnostic information. Replacement of the scatter plot displays in Figure 2 with Venn diagrams based on binary classification of cases according to single optimal thresholds would lead to a loss of this vital information.

Comment 10: P19 ln9 # 23, P20 ln 1 #5, and Figure 5. As these 29 cases are excluded, and the definitive statuses of these cases are not known, the analysis of these data separately adds nothing to the main body of data. Consequently, this section and Figure 5 should be eliminated, or placed in the Supplement.

Response: We agree that these components contribute little new information. Figure 5, its caption (p 42 line 8, deleted) and the associated text in Results p 19 line 16 (deleted) have accordingly been removed.