Neurochemical Dementia Diagnostics – Interlaboratory Variation of Analysis, Reference Ranges and Interpretations

Hansotto Reiber*, Peter Lange1 and Inga Zerr2

1Soc. for Advancement of Quality Assurance in Medical Laboratories, INSTAND, Düsseldorf, Germany
2Neurochemistry Laboratory, Department of Neurology, University Medical Center, Georg August University Göttingen, Germany

Abstract

Purpose: Dementia marker analysis requests the control of analytical reliability.

Method: 49 laboratories from nine European countries and USA participate in the first external quality assessment system [EQAS] for dementia marker analysis. Stabilized CSF samples are analyzed with a reference range-related evaluation and a differential diagnostic interpretation of combined parameters [total Tau protein, phospho Tau protein and Amyloid-ß-peptide Aß1-42].

Results: A) The large inter-laboratory variation is characterized by a survey-example with values in the decision range: the highest value was 10-fold higher than the lowest for Aß42 [69-771 pg/ml], 4-fold higher for Tau [315-1292 pg/ml and twofold for pTau [53-83 pg/ml]. With a success range of median ± 25% the fraction of outliers were up to 31% [Aß42] or 13-15% [Tau] and 3-11% for pTau in the N= 6 surveys. B) For evaluation [normal/pathological/border line] participants used a huge range of individual cut-off values: Tau [150-540, median 450 pg/ml], pTau [35-85, median 61 pg/ml] and Aß1-42 [205-600, median 500 pg/ml] with serious consequences for the differential diagnosis. C) In case of a sample with normal median values [e.g. Tau = 381 pg/ml and Aß= 748 pg/ml] 45% of participants regarded their values as pathological with a stunning interpretation of combined Tau and Aß1-42 data: 29% of the participants found this data combination compatible with an Alzheimer’s disease, 29% reported this as a normal sample, and 42 % regarded an interpretation as not possible.

Conclusions: Up to 31% outliers are a source of serious diagnostic errors. The unacceptable large variation of the laboratory own cut-off values leads to false negative and false positive diagnostic interpretations. This questions the practical relevance of dementia marker analysis. The calculation of mathematical formulas or ratios of the analytical parameters is not improving the discriminative sensitivity due to the error propagation in mathematical functions.

Keywords: Alzheimer’s disease; Creuztfeldt Jakob disease; CSF analysis; Dementia marker; External Quality assessment [EQAS]; Inter-laboratory accuracy; Aß1-42 amyloid peptide; Tau protein; phospho Tau protein; Reference values

Introduction

For differential diagnosis of dementia in neurological diseases, in particular Alzheimer’s disease [AD], a set of marker molecules in cerebrospinal fluid [CSF] was introduced and their relevance was investigated in many studies [1-4]. There is no doubt that the combined data for the mean values of increased total tau protein [Tau] and decreased Aß1-42 amyloid peptide [Aß42] allow a statistically significant discrimination of AD from non-AD patients. Additional analysis of hyper phosphorylated Tau protein [pTau] [1,5] may improve the differential diagnosis between different dementive processes [5-8]. Severe increase of Tau protein with decreased Aß42 may point to the rare Creutzfeldt Jakob disease [CJD] [6,7] to initiate further analysis with more specific parameters.

A large inter-laboratory variability of data was recognized by the international scientific community and lead to common efforts to reduce the pre-analytical problems with Aß peptide analysis [9] and to get reliable reference values for the parameters analyzed [1-3,5]. These efforts with controlled analytical protocols in well-trained expert groups [2,5,9] have shown the potential of this analysis. But with the spread of the analysis to many not specialized and less informed laboratories the daily practice for the analysis of the single patient became an important issue for the practicability and reliability of this neurochemical support of dementia diagnosis. The invention of the external quality control by independent institutions was an important step to help the clinical chemist to control the performance of his laboratory for accuracy and variability of the methods. Since 2010 the European Institution of quality assessment, INSTAND, Germany, offers a corresponding external quality assessment system [EQAS] for the neurochemical dementia marker analysis in the frame of its general survey for CSF analysis [10,11].

It is part of the concept of INSTAND to offer samples for analysis with data in a clinically relevant range, which allows also the reference range-related evaluation and for the combined set of data of these parameters interpretations relevant for differential diagnosis of dementia and other diseases of the nervous system. This approach represents a more demanding quality assessment compared to a sole certification of numerical values [10]. This demand of a general quality assessment led us to the extension of the survey, by requesting the laboratories own cut off values, their decision base for identification of a normal versus a pathological value. Additionally by multiple choice questions we offered the chance to improve the knowledge base in the

*Corresponding author: Hansotto Reiber, Soc. for Advancement of Quality Assurance in Medical Laboratories, INSTAND, Düsseldorf, Germany, Tel: 005511982785749; E-mail: ho@horeiber.de

Received March 20, 2014; Accepted April 22, 2014; Published May 20, 2014


Copyright: © 2014 Reiber H, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
single laboratory for differential diagnostic interpretations. By this concept we got a reliable description of the reality of the daily analytical practice in different laboratories, different from expert surveys [8] with controlled, common pre-analytical and analytical protocols. Due to the recent development of procedures for sample stabilization [12] we could avoid in the survey for distribution of samples the frequent pre-analytical problems.

Three years' experience with meanwhile 6 surveys and 50 participants from 9 European countries and USA gives an alarming result, which questions the reliability of this diagnostic approach in its actual performance.

The analysis of these marker proteins is not only a scientific but also an economic problem. Only slowly, with improved knowledge of the neurologists and psychiatrists, the indication for the request of this analysis was improved, as a basic contribution to cost reduction in the laboratories for these relatively expensive assays of different parameters with different relevance and qualities.

Methods

Sample preparation

Confectioning, preparation and stability control of CSF samples for the survey was performed by Peter Lange, Neurochemistry Laboratory, University Goettingen. Prof. Dr. Inga Zerr, co-advisor of the survey, head of Neurochemistry Laboratory surveyed the clinical aspects and diagnosis. Concepts for sample confection and selection of sample pairs are in the responsibility of Prof. Dr. Hansotto Reiber, as advisor in charge for performance of the CSF survey on behalf of INSTAND eV, Düsseldorf, Germany [www.instand-ev.de].

Normal or pathological CSF samples are collected in pools kept at 4°C. These samples originate from residual volumes remaining after clinically indicated extraction for routine analysis. In one survey [May 12] the sample was ventricular CSF obtained from catheter of a single patient. Samples were stabilized with sodium azid [12] to avoid pre-analytical problems. The samples are stable at room temperature for 4 weeks and freezing/ thawing did not influence the recovery of the initial concentration values. Aliquot samples, kept at 4°C, are distributed without freezing by normal mail. Samples have been controlled to be stable between sending and the deadline for reporting of results to INSTAND.

Data and interpretation protocol

Together with the samples the participants get a form [data protocol] with request for analytical data, their reference range-related evaluation, their reference ranges [cut off values] used, and a multiple choice set of questions for the differential diagnostic interpretation of the combined data set. Our summary of the results refer to the N=6 surveys on "Neurochemical dementia diagnosis" distributed by INSTAND between May 2011 and Nov 2013 [s. commentary/reports on www.instand-ev.de]. In this time interval the clinically oriented questionnaire for interpretation was several-fold improved. The success interval for certificates was determined [target value ± 25%]. Target value is the median of the group. Means of groups and SD were calculated from the subgroup after exclusion of the outliers [values outside the success interval].

Assays and statistics

The analytical assays used by the participating laboratories were from the same supplier [Innogenetics, Zwijnaarde, Belgium] with one exception in which a multianalyte system was used. This allowed a common evaluation of all participants, increasing from N=30 in the first of the surveys [May 2011] to N= 49 in the last survey [Nov 2013] integrated in this study.

Statistic treatment used standard procedures for calculation of means and standard deviation [SD] or corresponding coefficients of variation [CV=SD/mean x 100 in %].

Results

Inter-laboratory variation of analytical data

The inter-laboratory variation of analytical values is shown with median, coefficient of variation [CV%] and range of values as reported from N=42 participating laboratories (Table 1). This example [survey Nov. 2012] was selected, as the median values of the parameters are representative for the most relevant concentrations in the decision range between normal and pathological values in CSF. The range of concentration values reported by the participating laboratories in this survey is very large. The largest value was 10-fold higher than the lowest for Aß42 [69-771 pg/ml], 4-fold higher for Tau [315-1292 pg/ml] but less then twofold for pTau [53-83 pg/ml]. Nevertheless the success quotas were in an acceptable range referring to the success interval [median ± 25%].

The summary for the inter-laboratory variation for the CSF concentrations from all N= 6 surveys are shown in relation to the concentrations in Table 2. The coefficients of variation [CV] for the residual subgroups after elimination of the outliers are between 9.6% and 13.7% (Table 2). The success quota for Aß42 was between 69 and 92% and for tau protein between 85 and 87%. The best performance was found for pTau with 89-97%. The corresponding figures for the outliers with up to 31% [Aß1-42] were alarming.

Success interval for Tau, pTau and Aß42

As a common practice, the size of the success interval in a survey is determined after analysis of the data distribution in the group of participants. From the six surveys, so far, we got reliable coefficients of variation for the parameters Tau, pTau and Aß 1-42.

<table>
<thead>
<tr>
<th>Median</th>
<th>Range</th>
<th>CV %</th>
<th>N</th>
<th>Quota %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aß1-42</td>
<td>508</td>
<td>69-771</td>
<td>12</td>
<td>42</td>
</tr>
<tr>
<td>tTau</td>
<td>921</td>
<td>315-1292</td>
<td>13.3</td>
<td>42</td>
</tr>
<tr>
<td>pTau</td>
<td>181</td>
<td>66-53-83</td>
<td>9.6</td>
<td>34</td>
</tr>
</tbody>
</table>

Table 1: Inter-laboratory variation in the survey "Neurochemical dementia analysis" of INSTAND, Düsseldorf (survey Nov. 2012). This example was selected as the median concentration values (pg/ml) of the parameters are in the most relevant decision range between normal and pathological values in CSF. The coefficient of variation (CV), the number of participating laboratories and the success quota for the success interval (median ± 25%) are interpreted in the commentary of this survey (Nov 12, www.instand-ev.de).

<table>
<thead>
<tr>
<th>Conc.</th>
<th>CV %</th>
<th>Quota %</th>
<th>Conc.</th>
<th>CV %</th>
<th>Quota %</th>
<th>Conc.</th>
<th>CV %</th>
<th>Quota %</th>
</tr>
</thead>
<tbody>
<tr>
<td>242</td>
<td>13.7</td>
<td>69</td>
<td>&gt;220</td>
<td>11.7</td>
<td>87</td>
<td>37-53</td>
<td>11</td>
<td>90-95</td>
</tr>
<tr>
<td>&gt;500</td>
<td>12</td>
<td>73-88</td>
<td>460</td>
<td>13.6</td>
<td>85</td>
<td>66-68</td>
<td>9.6</td>
<td>89-97</td>
</tr>
<tr>
<td>&gt;800</td>
<td>13</td>
<td>75-92</td>
<td>920</td>
<td>13.3</td>
<td>86</td>
<td>74</td>
<td>10.7</td>
<td>97</td>
</tr>
</tbody>
</table>

Table 2: Summary of the inter-laboratory variation and success quota for the different concentration ranges (pg/ml) in the samples of the N= 6 surveys between 2011 and 2013. CV % refers to the mean of the groups after elimination of the outliers. Success quotas are the fraction of values from all participants which are in the success interval (median ± 25%).
In a Gaussian distribution the range of mean ± 2 SD or mean ± 2 CV includes 96% of the data of the whole group. Correspondingly, with a mean CV – 12.5% (Table 2) we calculated for Aß42, Tau and Phospho Tau the success interval as TV ± 25% [TV = target value= median, s. survey Nov 2011]. The corresponding quotas are compared in Table 2. A success interval of ± 30% would increase the quota only with about 3% but the interval of ± 20% would decrease the quota by 4-6%.

As a consequence, in the surveys “Neurochemical dementia diagnostics” we use for Aß 42, Tau and Phospho Tau the empirically founded success intervals of TV ± 25% (excluding 3-31% of the participants (Table2).

**Cut off values for reference range-related data evaluation**

The data of the cut off values used in the individual laboratory for evaluation of the analytical data are shown in Table 3. Obviously there is no consensus in the field. The upper values are the 3 to 4 fold of the lower reference values (Table 3).The consequences are unbearable as shown in Table 4: A tau protein concentration Tau = 381 ng/ml is regarded as normal by 43%, as borderline by 7% and as pathological by 50% of the participants (survey Nov 11)). A value of Aß1-42 = 508 ng/ml with an analytical data range of 69-771 pg/ml (Table 1) was regarded as normal by 35%, as border line by 30% and as pathological by 35% of the participants [survey Nov 12, Table 4]. Only in cases of extreme values the participants reached a higher correctness, but not 100% of correct evaluations: With a mean Aß42 value of 809 pg/ml and an analytical data range between the different laboratories of 420-1280 ng/ml [Nov 13, Table 4] only 70% correct evaluations were reported. For the mean Aß42 value of 242 pg/ml with an analytical data range of 136-393 pg/ml (May 12, Table 4) 97% of the values were reported correctly. The detailed variation ranges of evaluations are reported in the commentaries [www.instand-ev.de].

**Table 3: Cut off values (reference ranges) reported by the participating laboratories for their evaluation of analytical data as normal or pathological (example from last survey, Nov 13).** The median values (pg/ml) are representative for all surveys in the last 3 years. The huge range of the cut-off values used in the different laboratories is the source of false positive and false negative evaluations of the sample concentrations analyzed by the individual laboratory.

<table>
<thead>
<tr>
<th>Analyte</th>
<th>Range</th>
<th>Median</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aß1-42</td>
<td>205-600</td>
<td>500</td>
<td>42</td>
</tr>
<tr>
<td>Total Tau</td>
<td>150-540</td>
<td>450</td>
<td>40</td>
</tr>
<tr>
<td>Phospho Tau</td>
<td>35-85</td>
<td>61</td>
<td>39</td>
</tr>
</tbody>
</table>

**Table 4: Evaluations and interpretations of analytical values by the individual laboratories. These interpretations are based on the largely varying cut off values reported by the participants (Table 3). The data of the five surveys [Nov11-Nov.13] are reported and explained in the commentaries documented in the website of INSTAND [www.instand-ev.de]. For different combinations of the analytes Aß1-42 amyloid peptide and total Tau protein concentrations we show the differences reported by the individual laboratories according to their individual cut off values [reference range-related evaluation].**

<table>
<thead>
<tr>
<th>Analyte</th>
<th>Concentration (pg/ml)</th>
<th>Cut-off related evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aß 42</td>
<td>748</td>
<td>440-2024 25/30</td>
</tr>
<tr>
<td></td>
<td>381</td>
<td>99.8-654 13/30 15/30</td>
</tr>
<tr>
<td></td>
<td>242</td>
<td>136-393 38/39</td>
</tr>
<tr>
<td></td>
<td>459</td>
<td>326-650 14/36 25/36</td>
</tr>
<tr>
<td></td>
<td>508</td>
<td>69-771 14/40 14/40 12/40</td>
</tr>
<tr>
<td></td>
<td>921</td>
<td>315-1292 40/40</td>
</tr>
<tr>
<td></td>
<td>978</td>
<td>696-1311 40/47 1/47</td>
</tr>
<tr>
<td></td>
<td>3260</td>
<td>1097-5280 41/48</td>
</tr>
<tr>
<td></td>
<td>809</td>
<td>420-1280 33/47 1/47</td>
</tr>
<tr>
<td></td>
<td>222</td>
<td>127.5-345 33/47 1/47</td>
</tr>
</tbody>
</table>

**Table 5: Data combinations of Aß42 and Tau in the CSF surveys Nov11-Nov13 together with the differential diagnostic interpretation by the participating laboratories according to the multiple choice criteria: Normal dementia marker, data compatible with AD, data compatible with CJD, data can not be interpreted. Detailed results and explanations are shown in the commentary of the surveys [www.instand-ev.de].**

**Interpretations for differential diagnosis of dementia**

The main target of this analysis is the contribution to differential diagnosis of dementia. With the large analytical imprecision and the missing of reliable, common reference ranges it is not astonishing what we got as a most confusing result from the surveys (Table 5). A definitely normal combination of data [Nov 11] was interpreted as an Alzheimer’s disease by 8/28 laboratories and only 8/ 28 found these combined results as normal. In case of definitely increased Tau protein in CSF and definitely decreased Aß42 values [Nov 12] 29/36 laboratories were sure to have the pattern of an Alzheimer’s disease, 3/26 kept it as indicative for a Creutzfeldt-Jacob disease and only 4/36 laboratories reported the most probable result as a not interpretable combination. Vice versa, in case of a definite Alzheimer’s disease pattern with increased Tau and definitely decreased Aß1-42 [May 12] 6/35 laboratories could not make an interpretation.

These data do not allow regarding this analytical approach as a reliable neurochemical support for the diagnosis of neurological diseases with a dementia.

**Discussion**

**Sample quality**

The controlled stability of the CSF samples distributed by INSTAND for the CSF survey excludes pre-analytical problems [9]. The homogeneity of the distributed sample aliquots is systematically controlled, to guaranty the reliability of the reported inter-laboratory variations. All CSF samples used for the survey either pooled or from a single patient, were routinely analyzed for the complete set of data [13] like cell count, hemoglobin, albumin and immunoglobulin quotients or lactate to control the quality of the material used.

**Analytical inter-laboratory variation**

The precautions in sample preparation allow the conclusion that the large inter-laboratory variation in the survey is a consequence of the bad performance of the assays. The 10-fold higher value of the largest concentration compared to the lowest value from the participants for Aß42 or a 4-fold difference for Tau protein (Table 1) need urgent action of the assay supplier. It makes this analysis useless if up to 31% of the values are outside the success interval in this case with median ± 25%. These deficits in accuracy of the values in the individual laboratory could be controlled by reliable control samples with concentrations in
The decision range [cut off]. An improvement of the assays would be possible as shown by the pTau results with less than twofold variation [53-83 pg/ml, Table 1] and only 3-11% outliers corresponding to success quota of 89-97% (Table 2). It remains to hope that the new assays on the market [9,14] avoid these deficits by more stable calibrator samples and better test robustness together with suitable control samples.

Reference ranges

The cut off values between normal and pathological values of the biomarkers still after 6 surveys the participants were far from a consensus (Table 3). The threefold larger upper value compared to the lower cut-off value led to the stunning misinterpretations reported in Table 4. Based on the data in several publications [1-3,15] we get the average values of the means and standard deviations in Table 6. The usual calculation of the reference ranges as mean ± 2 SD includes 96% of the normal controls [cut off values in Table 6]. These data would mean that normal Tau protein values are found up to 510 pg/ml and normal Aβ1-42 values are found as low as 310 pg/ml. If these values would be taken as cut-off values, there would be many data of Alzheimer’s patients in the normal range. The biological range of Alzheimer’s data calculated as mean ± 2 SD (Table 6) as an average of the data from the literature [1-3,15] has a strong overlap with the range of normal controls. A cut off value for the reference range of tau protein with <510 pg/ml is in the lower part of the range for AD patients [50-1350 pg/ml]. As a consequence we get 30-40% of false negative interpretations for AD patients, but for any lower cut-off value we would get false positive interpretations of normal controls. The same is the case for Aβ1-42 with a possible cut-off value of 310 pg/ml. The overlap of the ranges would be still larger if the inter-laboratory variation would be considered. These considerations show the combined uncertainties for the evaluation of the data from an individual patient and explain why false positive and false negative evaluations like those for the survey in Table 4 must be unavoidable for the individual patient in the daily practice.

Tau protein in ventricular CSF, the rostro-caudal concentration gradient

The CSF sample in survey May 2012 was ventricular CSF collected from catheter of an individual patient without a dementia. As shown earlier [16,17] normal Tau protein values in ventricular CSF are 1.5 fold higher than in lumbar CSF. This means that the tau value of 456 pg/ml in ventricular CSF (Table 2) would correspond with 300 pg/ml in lumbar CSF of this patient. This would fit a normal CSF value for Tau protein. The unquestionably decreased Aβ1-42 value in this sample may have been a consequence of the collection process with a catheter. Due to a rostro-caudal, decreasing concentration gradient of proteins from neurons and glial cells [16,17] the concentration in lumbar CSF must increase with an increasing extraction volume in a lumbar puncture.

This dependency of the reference range from CSF extraction volume has also to be taken into account for the dementia marker proteins.

Combined data interpretations for differential diagnosis of dementia

Due to the large variation of concentration values and the missing consensus about reference ranges with the many false pathological and false normal evaluations also the differential diagnostic relevant interpretation of combined data must largely fail. An assay in which a normal CSF sample is interpreted by 29% of the laboratories as compatible with the diagnosis of an AD (Table 5) is not suitable for support of the diagnosis in daily routine. As a solution of this unsatisfactory situation several authors tried to extend the spectrum of dementia marker proteins [15].

Other parameters and evaluation ratios

After an initial trial to integrate the peptide Aβ1-40 and the Aβ1-42/Aβ1-40 ratio in the survey, this analysis was omitted, as the participating group was too small with a too large analytical variation. Additionally a consensus about the dimension of the ratio has been missing. In general the calculation of Aβ42/Aβ40 ratio or Aβ42/Tau ratio [3,15] and other formulas like Innoteq Amyloid Tau Index [4] for a combined evaluation of the biomarkers are not useful due to the error propagation in mathematical functions which leads in case of biochemically or analytically uncoupled parameters to a much larger imprecision. The CV of a quotient, i.e., the division of two uncoupled parameters, is the sum of the CVs of the two individual parameters.

Actually, as a consequence dementia disorders remain primarily a clinical diagnosis without support by the dementia marker analysis. It will be the most demanding challenge to work for a consensus about cut off values as a reasonable common base for an acceptable inter-laboratory variation of patient data. This needs better assays with a smaller analytical variation and reasonable algorithms for the combination with other CSF data and clinical information.

Actually the positive contribution of CSF analysis to the differential diagnosis comes from the analysis of a complete spectrum as described [11,13] to get the differentiation between inflammatory, non-inflammatory and neurodegenerative neurological diseases [15].

Acknowledgement

We thank Prof. M. Otto, Ulm, for many helpful discussions in the planning of the dementia biomarker survey and Prof. Dr. P. Lewczuk, Würzburg, for cooperation and supply of samples in the initial phase during the establishment of the survey.

References


