

---

## Research Article

---

# Development and Advanced Validation of an Optimized Method for the Quantitation of A $\beta$ <sub>42</sub> in Human Cerebrospinal Fluid

Valerie C. Cullen,<sup>1</sup> Ross A. Fredenburg,<sup>1</sup> Cindy Evans,<sup>2</sup> Phyllis R. Conliffe,<sup>2</sup> and Michael E. Solomon<sup>1,3</sup>

Received 27 December 2011; accepted 5 April 2012; published online 3 May 2012

**Abstract.** Cerebrospinal fluid (CSF) biomarkers have been extensively utilized in the diagnosis of Alzheimer's disease (AD) and characterization of progression. One important CSF biomarker is the amyloid beta 42 (A $\beta$ <sub>42</sub>) peptide, a key player in AD pathogenesis. The INNOTEST® A $\beta$ <sub>42</sub> ELISA kit has been widely used but an advanced level of method development and validation has not been reported. To support a clinical trial in AD, we successfully completed a Good Laboratory Practices (GLP)-level validation of the method to establish the parameters of precision, accuracy, parallelism, selectivity, specificity, and linearity of dilution of the assay in CSF matrix, as well as CSF storage stability. Several modifications were required to optimize the assay and ensure consistent results in a clinical-trial setting. These included the use of additional calibrators, an adjusted standard curve range, a minimum required dilution (MRD) of CSF by 6-fold to avoid matrix interference and mitigation of analyte adsorption to labware by the addition of Tween-20. The optimized method displayed a quantitative range of 375–4,500 pg/mL. The inter-assay precision was  $\leq 12.1$  % CV and the inter-assay relative accuracy was  $\leq 10.9$  % absolute bias, bringing the total error of the assay to  $\leq 23$  %. The intra-assay precision of the assay at the high validation standard and below was  $\leq 5.5$  % CV; this enables sensitive detection of biomarker changes across a therapeutic regime. The INNOTEST® A $\beta$ <sub>42</sub> ELISA kit, modified as reported here, may be appropriate for many applications, including regulatory agency acceptable clinical diagnosis and pharmacodynamic assessment.

**KEY WORDS:** AB42; Alzheimer; biomarker; enzyme-linked immunosorbent assay; method validation.

## INTRODUCTION

Alzheimer's disease (AD) affects millions worldwide, with rising incidence as populations age. Despite extensive efforts across industry and academia, an effective strategy for disease modification remains elusive. Amongst the main concerns has been how to predict who will succumb to AD and how to monitor clinical improvement during therapeutic intervention. To this end, AD biomarkers have received much attention, including CSF amyloid beta 42 (A $\beta$ <sub>42</sub>), which is lower in AD patients than in normal controls (1–6). Low CSF A $\beta$ <sub>42</sub> levels also predict conversion from mild cognitive impairment (MCI) to AD (7–12).

Most A $\beta$ <sub>42</sub> assays rely on immunochemical detection, whether conventional ELISAs, electrochemiluminescence-based ELISAs or bead-based methods. One of the most widely used assays is the INNOTEST® ELISA from Innogenetics NV which has been cited in numerous publications from many study centers (3–5,13–16).

Published studies using this kit invariably conduct the assay on neat CSF according to the manufacturer's instructions. To our knowledge, no group has undertaken an independent and comprehensive method development and validation study for this assay.

In the present study, we first undertook an extensive method development study of the INNOTEST® A $\beta$ <sub>42</sub> ELISA, creating a number of important modifications to the assay. These were: (1) establishment of the minimum required dilution (MRD) for CSF samples, which upon empirical testing in both AD and normal human CSF was a dilution of one in six, (2) changes in the calibration curve design to better bracket A $\beta$ <sub>42</sub> concentration ranges observed in clinical samples, and (3) introduction of quality control (QC) samples composed of human CSF from pre-mortem AD donors.

We also systematically examined A $\beta$ <sub>42</sub> recovery from storage and assay vials. Although polypropylene is considered to be the least problematic polymer for adherence of hydrophobic peptides (17–19), it is known that A $\beta$ <sub>42</sub> can variably adsorb to polypropylene vials made by different

---

Valerie C. Cullen and Ross A. Fredenburg contributed equally to the work described.

**Electronic supplementary material** The online version of this article (doi:10.1208/s12248-012-9360-7) contains supplementary material, which is available to authorized users.

<sup>1</sup> Link Medicine Corporation, 161 First Street, Cambridge, Massachusetts 02142, USA.

<sup>2</sup> Tandem Labs Biotechnology Services (A Division of Labcorp), 13112 Evening Creek Drive South, San Diego, California 92128, USA.

<sup>3</sup> To whom correspondence should be addressed. (e-mail: mesolomon@yahoo.com)

manufacturers (20). Since such adsorbance would impair assay precision and accuracy, we tested different polypropylene vial types for A $\beta$ <sub>42</sub> recovery after freeze/thaw, and we established a mitigation strategy in the event of A $\beta$ <sub>42</sub> adsorbance.

Once these modifications were established, we then executed a full method validation in which the assay was characterized in terms of precision, accuracy<sup>1</sup>, selectivity, specificity, parallelism, linearity of dilution, and stability. The guidelines for biomarker assay validation set out by Lee *et al.* (21) were used as guiding principles. As presented herein, the assay performs very well with excellent precision, accuracy and reliability.

## MATERIALS AND METHODS

### Materials

INNOTEST®  $\beta$ -AMYLOID<sub>(1-42)</sub> assay kit was obtained from Innogenetics (Ghent, Belgium) and utilizes a 21F12/3D6 sandwich pair and a A $\beta$ <sub>1-42</sub> peptide solution as calibration standard. Single-subject lots of normal and AD donor CSF were obtained from Precision Med, Inc. (San Diego, CA). Polypropylene assay tubes were obtained from Eppendorf (Hamburg, Germany; LoBind 1.5 mL #0030 108 116) and Thermo Fisher Scientific (Waltham, MA; 1 mL Nunc cryotubes #375353).

### Methods

#### Method Development

**Immunoassay Procedure.** For initial method development, the assay was performed per manufacturer directions. See [ELECTRONIC SUPPLEMENTARY MATERIALS](#) for details. Assay modifications were made throughout the method development process; final assay conditions are described in “[Method Validation](#)” below.

**MRD Determination.** Donor CSF samples were diluted in INNOTEST® Sample Diluent in Eppendorf LoBind tubes prior to loading of the set-up plate. The assay procedure was followed as above from that point forward. The dilution factor beyond which the back-calculated analyte concentration values remained consistent was deemed the MRD. All subsequent development assays were performed with an upfront dilution by the MRD value determined in these series of experiments (MRD=6).

**A $\beta$ <sub>42</sub> Adsorption Testing.** Three individual lots of donor CSF were thawed and tested to provide a “baseline” A $\beta$ <sub>42</sub> value. 250  $\mu$ L of each lot was transferred to a separate Nunc vial, refrozen overnight at  $-70^{\circ}$ C, thawed and re-tested. In addition, pooled donor CSF designated low (LQC) and mid-QC (MQC) and donor CSF spiked with A $\beta$ <sub>42</sub> calibrator (designated high QC (HQC)) were prepared in LoBind tubes and aliquotted

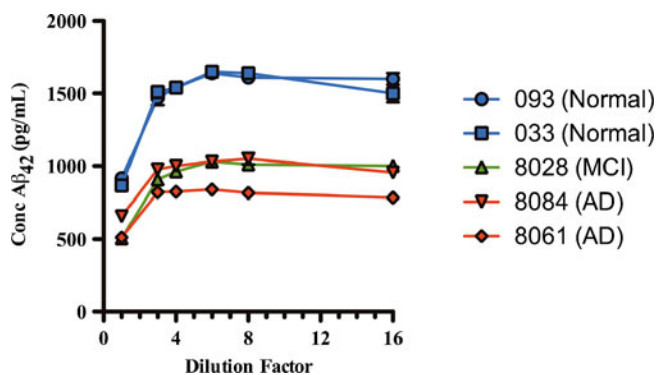
for  $-70^{\circ}$ C storage. For each QC level tested, a single aliquot of CSF was thawed and tested to provide a “baseline” A $\beta$ <sub>42</sub> value. Prior to re-freezing, 50  $\mu$ L of each lot was transferred to a separate LoBind tube, refrozen overnight at  $-70^{\circ}$ C, thawed and re-tested. A $\beta$ <sub>42</sub> levels before and after freeze-thaw were compared to determine loss of analyte in each tube type. See Supplementary Fig. 1a, b for an illustration.

**Adsorption Mitigation.** A previously unthawed vial of each of three AD donor CSF lots was thawed and aliquotted as follows: one 50- $\mu$ L volume was transferred to a LoBind tube and two 100- $\mu$ L volumes were transferred to two Nunc vials. All aliquots were frozen overnight at  $-70^{\circ}$ C. Upon thawing, the LoBind tube and one of the Nunc vials were loaded onto the assay plate with no treatment aside from dilution to MRD. The other Nunc vial was spiked with 10  $\mu$ L 2.2 % Tween-20 (in Sample Diluent), vortexed, incubated at room temperature for 30 min, and vortexed again before assaying. The resulting A $\beta$ <sub>42</sub> concentration values were compared to assess the ability of 0.2 % Tween-20 to mitigate A $\beta$ <sub>42</sub> adsorption. See Supplementary Fig. 2 for an illustration.

#### Method Validation

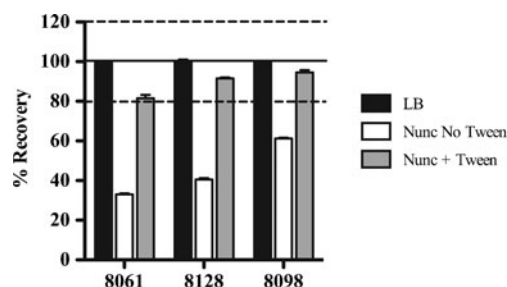
**Preparation of Calibration Curve.** The calibration standard (CS) material (A $\beta$ <sub>1-42</sub> Peptide; 91,163 pg/ml) from the kit was used to prepare ten standards at the nominal concentrations of 40.0, 50.0, 62.5, 125, 175, 250, 375, 500, 750 and 1,000 pg/mL in sample diluent with 0.033 % Tween-20. Unspiked sample diluent with 0.033 % Tween-20 was used as a buffer blank. The points above and below the target upper limit of quantitation (ULOQ) and lower limit of quantitation (LLOQ) were used as anchor points in the standard curve. Each standard was assayed as 1 duplicate set ( $n=2$ ).

**Standard Curve Evaluation.** Six standard curves were evaluated in six independent assays. Standard curves were fitted using 4-Parameter Logistic regression and (1/Y) weighting using Watson LIMS™ v. 7.03; the fit was selected based



**Fig. 1.** MRD determination in 5 individual donor CSF lots. CSF from normal elderly individuals (*blue lines*) and elderly individuals diagnosed with mild cognitive impairment (*MCI, green line*) or Alzheimer’s disease (*AD, red lines*) was diluted in sample diluent to various levels prior to loading into the assay plate. The back-calculated A $\beta$ <sub>42</sub> concentration for each lot at each dilution level was plotted against the applied dilution factor. Error bars describe the range around the mean of two replicates

<sup>1</sup> Due to the ubiquitous presence of endogenous A $\beta$ <sub>42</sub> in normal human CSF, blank native matrix was not available for these studies and the accuracy determinations generated were by definition relative. Due to the intended use of the assay, namely to measure changes in biomarker concentration across a treatment regime, absolute accuracy was not necessary and therefore CSF validation and quality control samples were created to best evaluate precision in the intended matrix.



**Fig. 2.** Mitigation of  $A\beta_{42}$  adsorption to polypropylene via Tween-20 addition. AD donor CSF lots were tested after an overnight incubation at  $-70^{\circ}\text{C}$  in Eppendorf LoBind tubes (black bars), after an overnight incubation at  $-70^{\circ}\text{C}$  in 1.0 mL Nunc cryovials but prior to Tween-20 addition (white bars), and after an overnight incubation at  $-70^{\circ}\text{C}$  in 1.0 mL Nunc cryovials and following the addition of Tween-20 to achieve a final Tween concentration of 0.2 % (see “Materials and Methods”) (gray bars). For each CSF lot, the  $A\beta_{42}$  concentration determined after LoBind incubation was set to 100 % in order to calculate the percent recovery of the other treatment methods. Error bars describe the range around the mean of two replicates

on analysis of over 20 method development runs. Acceptance was based on the following precision and accuracy criteria for each non-anchor CS level: back-calculated concentrations  $\leq 25.0$  % CV ( $\leq 30.0$  % at LLOQ and ULOQ) and within  $\pm 25.0$  % Bias ( $\pm 30.0$  % at LLOQ and ULOQ). Individual CS levels could be deactivated according to the above criteria, as long as at least 6 non-anchor points remained and no consecutive CS levels were deactivated.

**Preparation of Validation Samples.** Five levels of validation samples (VS) were prepared and used for accuracy and precision studies. VS2 and VS3 were created by pooling pre-screened AD donor CSF samples to obtain approximate endogenous levels of 870 and 2,200 pg/mL, respectively. VS2 and VS3 pools were aliquotted and frozen for use for the remainder of the validation as LQC and MQC (see below). Additionally, VS2 was aliquotted into vials to create VS1 (LLOQ; 375 pg/mL) by dilution in Sample Diluent and VS4 (3,600 pg/mL) and VS5 (ULOQ, 4,500 pg/mL) by spiking with  $A\beta_{42}$  calibrator on the day of assay. Additionally, two individual incurred samples (IS) were analyzed during the core runs and mean endogenous analyte levels were qualified for use in all stability assays. Each VS and IS sample was treated with 2.2 % Tween-20 at 10 % of total volume for at least 30 min and then diluted 6-fold in Sample Diluent for a final Tween-20 concentration of 0.033 %. Unless otherwise noted, all CSF samples were treated in this manner in the remaining validation experiments. A dilution factor of 6.6 was applied in Watson LIMS™ to accommodate MRD and Tween-20 addition to each sample.

**Inter- and Intra-Assay Accuracy and Precision.** Six accuracy and precision runs were performed over 3 days by two analysts, with VS1–VS5 samples in five sets of duplicate wells ( $n=10$  wells).

The target acceptance criteria for inter- and intra-assay accuracy and precision was  $\leq 25.0$  % CV or absolute bias for VS2, VS3, and VS4;  $\leq 30.0$  % CV or absolute bias for VS1 (LLOQ) and VS5 (ULOQ). Mean analytical recovery calculations for all VS levels were based on the established  $A\beta_{42}$  concentrations defined above.

For subsequent validation runs, VS2, VS3, and VS4 were assayed in duplicate on all plates as LQC, MQC, and HQC, respectively. Examining the six accuracy and precision runs, we elected to amend the validation plan to set the QC and CS acceptance criteria to  $\leq 20$  % for CV and absolute bias for all samples except for the calibrators at LLOQ and ULOQ for which the acceptance was set to  $\leq 25$  %. For assay data to be accepted, at least six non-anchor CS levels and at least four of the six QC samples (with at least one acceptable QC at each of the LQC, MQC, and HQC levels) on each plate had to meet these adjusted criteria. Mean analytical recoveries for LQC and MQC were based on established endogenous  $A\beta_{42}$  concentrations determined for VS2 and VS3 during the core runs: 797 pg/mL for VS2 (LQC) and 2,161 pg/mL for VS3 (MQC). HQC continued to be prepared on the day of assay by spiking 2,730 pg/mL of  $A\beta_{42}$  calibration standard into VS (LQC); the established  $A\beta_{42}$  concentration used to calculate mean analytical recovery for HQC was therefore 3,527 pg/mL (2,730+797).

**Selectivity.** Ten individual AD donor CSF samples were assayed in duplicate both untreated and spiked with  $A\beta_{42}$  calibrator solution at an approximate MQC concentration (2,200 pg/mL). Endogenous values determined for the untreated sample were subtracted from that of the spiked sample and recovery was calculated as a percentage of the nominal spike level (2,200 pg/mL). The target criterion for recovery was  $\pm 20$  % bias from nominal for 80 % of samples.

**Dilutional Linearity/Hook Effect.** Dilutional linearity and hook effect was evaluated by spiking with  $A\beta_{42}$  calibrator solution to a combined level of 9,000 pg/mL into the LQC CSF solution and then diluting 2-, 5-, 10-, 25-, 50-, and 100-fold in sample diluent with 0.033 % Tween-20. Precision and accuracy of each sample with an established concentration falling between LLOQ and ULOQ were subjected the acceptance criteria defined after the core runs. Dilutional samples with established concentrations above the ULOQ were expected to read above ULOQ, indicating the absence of a hook effect within the limits of the concentrations tested.

**Parallelism.** Parallelism was assessed using three individual AD donor CSF samples with varying amounts of endogenous  $A\beta_{42}$ . Samples were tested neat and diluted to various levels in Sample Diluent; for this experiment, the assay MRD was not applied to the samples but Tween-20 was added to the neat solutions prior to diluting them. Parallelism was evaluated by determining the mean back-calculated concentration of all samples diluted beyond the point at which matrix interference is absent as described in Ref. (21); precision at  $\leq 20$  % CV of the mean described above was considered an indication of parallelism.

**Specificity and Interference from Hemolysis.** The effect of  $A\beta_{40}$  on the quantification of  $A\beta_{42}$  was evaluated using LQC and HQC CSF solutions analyzed in the presence of either  $A\beta_{40}$  at 20 %, 50 % and 100 % of the assay ULOQ or hemolyzed blood spiked at 1 in 10, 1 in 100 and 1 in 1,000 (10, 1, 0.1 % blood contamination, respectively). Acceptance was based on the accuracy and precision criteria established in the core runs.

**Stability.** High and Low CSF QCs in LoBind tubes and two individual CSF samples (IS1 and IS2) in Nunc tubes (three duplicate sets at each level) were used for stability evaluation. Two thirds of the stability QCs at each level must have met the acceptance criteria for %bias and %CV as defined after the core runs. If a stability time-point failed to yield acceptable results, then stability was deemed confirmed up to the previous time-point tested.

The above samples were tested for bench-top (room temperature), refrigerated (on ice), and long-term frozen (−70°C) stability, as well as freeze/thaw stability (IS only). Bench-top and refrigerated stability was assessed at 4–6 and 16–24 h, and both prior to and after Tween-20 addition. Freeze/thaw stability of IS CSF after Tween-20 addition was assessed after one, two, and three cycles. Long-term stability was assessed at 1 and 3 months for QC and IS samples prior to Tween-20 addition and in IS samples after Tween addition. Each sample was subjected to MRD and Tween-20 treatment (except when Tween had been added prior to stability assessment). HQC samples tested for long-term frozen stability were freshly spiked from LQC with A $\beta$ <sub>42</sub> calibrator solution immediately prior to analysis, and therefore the samples held for long-term stability all contained the LQC A $\beta$ <sub>42</sub> concentration.

**Compliance Statement.** The validation experiments described herein were conducted in a GLP-compliant environment with the Tandem Labs Quality Assurance Unit overseeing the laboratory's compliance program and assuring the quality and integrity of the validation test data generated. Unless otherwise stated, all procedures were performed in accordance with Tandem's internal standard operating plans and the TCAS10-082 Validation Plan.

## RESULTS

### Method Development

#### Validation Feasibility Testing

INNOTEST® ELISA performance was evaluated in a series of experiments that comprehensively simulated an advanced method validation (i.e., addressing all components of assay performance required by regulatory agencies) (21). The INNOTEST® kit is based on a set of A $\beta$ <sub>42</sub> synthetic peptide calibrators in assay buffer spanning 125 to 2,000 pg/mL, but contains no QC samples. We modified the calibration curve to include additional points, better defining the optimal fit to a nonlinear regression algorithm. The range of the curve was also adjusted as method development progressed to best situate QC samples in relation to the anticipated reference range. To provide suitable material for VS and QC samples, for establishment of the anticipated reference range, and for selectivity testing, incurred CSF samples were obtained both from ongoing clinical studies and from a commercial source. The results of most of these experiments are recapitulated in “Method Validation”, but two important assay modifications are described in detail below.

#### Establishment of the MRD

CSF samples from two AD, one MCI, and two normal individuals were tested at various dilution levels to determine the impact of matrix on the measurement of endogenous A $\beta$ <sub>42</sub>. Figure 1 illustrates that neat CSF generated A $\beta$ <sub>42</sub> concentration values that were on average 40–50 % lower than those reported for the same CSF samples after dilution. This apparent matrix interference was consistent across various analyte levels and was similar in normal, MCI and AD samples. Although dilution of CSF to 50 or 33 % (data not shown) relieved much of the apparent interference, dilution to 16.7 % CSF or lower was required to eliminate the effect. The back-calculated A $\beta$ <sub>42</sub> concentrations at 6-, 8-, and 16-fold dilutions of CSF were very similar to each other (CV of the mean of the three values, <10 %), indicating that parallelism exists between the calibration curve and the endogenous analyte upon appropriate dilution of the matrix. For these reasons, the MRD of the assay was determined to be six; all CSF samples and controls were diluted 6-fold for all subsequent experiments.

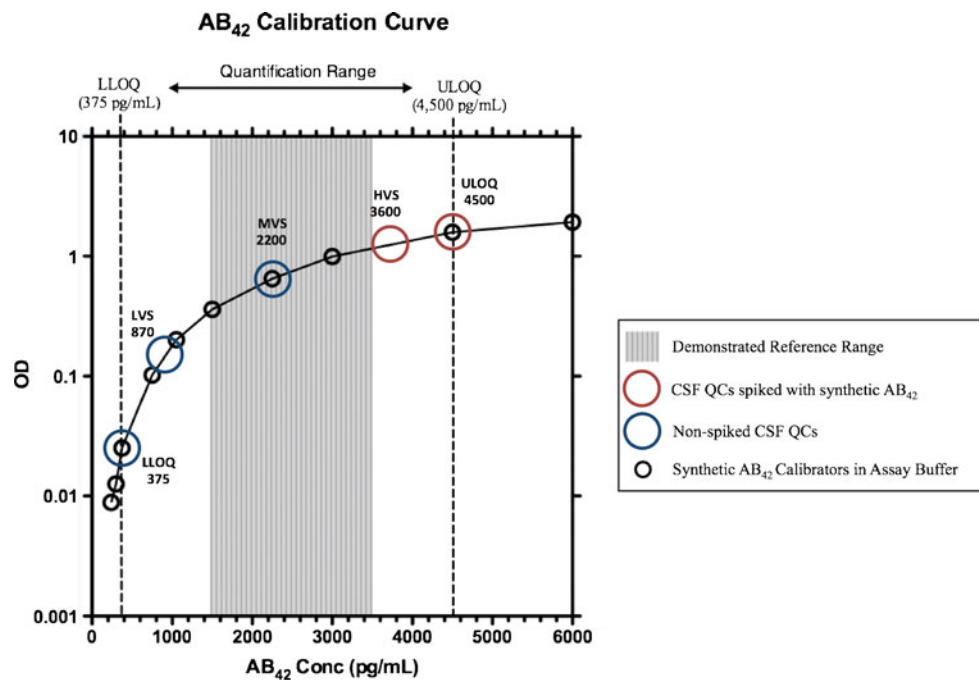
#### Detergent Mitigation of A $\beta$ <sub>42</sub> Adsorption to Polypropylene Vials

As discussed earlier, the recovery of A $\beta$ <sub>42</sub> from polymer vials during handling and quantification has been a recurring concern (18,20). As we had selected Nunc brand polypropylene cryovials (#375353) for the storage of clinical CSF samples and Eppendorf Protein LoBind tubes (#0030 108.116) for assay execution, we first evaluated both tube types for A $\beta$ <sub>42</sub> recovery after freeze–thaw as described in “Methods” and Supplementary Fig. 1a, b. These experiments revealed that A $\beta$ <sub>42</sub> recovery from LoBind tubes after a single freeze–thaw cycle is 91–95 % of the expected value, but that a significant and variable proportion (30–70 %) of A $\beta$ <sub>42</sub> is lost after a single cycle in Nunc cryovials (Supplementary Fig. 3). The >90 % recovery in LoBind tubes indicated that the low recoveries observed in Nunc vials was due to adsorption of A $\beta$ <sub>42</sub> to tube walls (as opposed to degradation or aggregation). To investigate adsorption mitigation, Tween-20 addition was tested as described in “Methods” and Supplementary Fig. 2. Figure 2 demonstrates that Tween-20 effectively recovers the proportion of CSF A $\beta$ <sub>42</sub> that would otherwise remain inaccessible in Nunc tubes. This is consistent with a report by McCush *et al.* (22) and provides an *ex post facto* adsorption mitigation strategy for CSF that was not treated with detergent prior to storage. The presence of Tween-20 at the final concentration of 0.033 % was innocuous to the performance of the calibration curve and to A $\beta$ <sub>42</sub> quantitation (Supplementary Fig. 4). The assay was therefore modified to include a 10 % (v/v) spike of 2.2 % Tween-20 in assay buffer to all CSF samples in method validation and any subsequent clinical analysis.

#### Summary of Assay Modifications After Method Development

Figure 3 illustrates the specifications of the modified INNOTEST® A $\beta$ <sub>42</sub> ELISA at the conclusion of method development. An approximate clinical reference range of 1,500 to 3,500 pg/mL was established by analyzing multiple





**Fig. 3.** Calibration curve and VS samples for method validation. The final calibration series consisted of synthetic  $A\beta_{42}$  peptide (small circles) in buffer at well concentrations of 1,000, 750, 500, 375, 250, 175, 125, 62.5, 50, and 40 pg/mL; anchor points were at 1,000, 50, and 40 pg/mL. The LLOQ and ULOQ of the method (dashed lines) were defined by the calibrators at well  $A\beta_{42}$  concentrations of 62.5 and 750 pg/mL, respectively. Five VS levels (large circles) were employed in the core runs and consisted of non-spiked AD donor CSF (blue circles) or CSF spiked with  $A\beta_{42}$  calibrator (red circles). The anticipated reference range determined from a limited number of incurred CSF samples is shown as a gray-shaded region

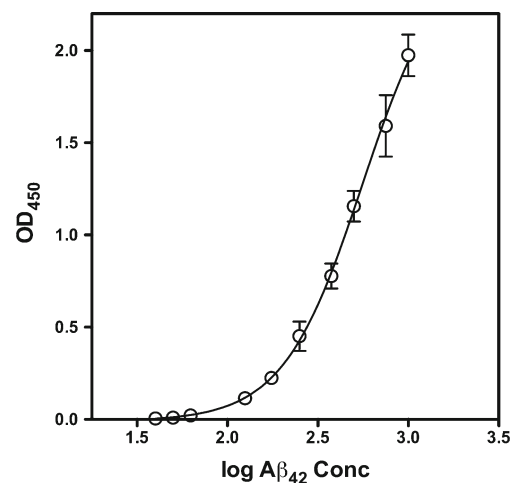
CSF samples from early AD patients both at an MRD of six and treated with Tween-20 (data not shown). The calibration range of the original assay extended to 2,000 pg/mL; since the incurred CSF samples were subjected to MRD and the buffer-based synthetic calibrators were not, the range of quantification was effectively magnified 6-fold (to 12,000 pg/mL) with the original calibration curve. To better span the expected concentration range of samples, the highest calibrator was reduced from 2,000 to 1,000 pg/mL per well (6,000 pg/mL after adjustment for MRD). A 10-point calibration curve was designed to place 7 points in the quantification range and include three anchor points. The LLOQ of the assay was established at 375 pg/mL (62.5 pg/mL calibrator in well) and the ULOQ was set at 4,500 pg/mL (750 pg/mL calibrator in well). Five VS samples were selected for the core runs of method validation as defined in “Methods” and Fig. 3. The two non-spiked VS samples (VS2 and VS3) were created by pooling individual incurred CSF samples of established  $A\beta_{42}$  concentration. These pools were qualified in two separate development runs to assign an endogenous value for calculation of the spike level for VS4 (high validation standard) and VS5 (ULOQ) as well as the dilution factor for creation of VS1 (LLOQ).

## Method Validation

### Calibration Curve Performance

Six validation runs were performed over 3 days by two analysts to establish the performance of the calibration curve.

Each run passed pre-defined acceptance criteria as specified in “Methods.” The precision of the interpolated values of all calibrators in the quantitative assay range was  $\leq 3.8\%$  CV. The accuracy was also excellent, with  $\leq 4.1\%$  absolute bias. Each calibrator series fit well to the four-parameter logistic algorithm with 1/Y weighting (see Fig. 4).



**Fig. 4.** Calibration curve and goodness of fit to 4-PL. The logarithm of the mean OD signal over all six core runs is shown for all ten  $A\beta_{42}$  calibrators. The line represents the best fit of the data to a four-parameter logistic function with 1/Y weighting. Error bars represent  $\pm 1$  standard deviation around the mean of six independent experiments

### Precision and Accuracy

Over six runs performed on three separate days by two analysts, all VS levels passed the *a priori* acceptance criterion of total error<sup>2</sup> (TE; %CV+[%bias]) $\leq$ 30 %. In fact, the highest value of TE was 18.4 % (observed in VS5/ULOQ). In Table I, the intra-assay precision and accuracy ranges at each VS level as well as the values for inter-assay precision and accuracy are displayed. Within a plate, the precision values range between 0.6 and 8.4 %CV, while between plates, the values fall between 4.2 and 12.1 %CV. All VS levels performed equally well whether they were constructed by dilution of a CSF pool (VS1), by using endogenous CSF pools (VS2 and VS3) or by spiking a CSF pool with calibrator (VS4 and VS5). The accuracy of the assessments varied to a greater degree than did the precision, ranging from -22.1 to 12.9 % bias (both extremes occurred at VS5/ULOQ), but all bias values fell within the pre-established acceptance criterion of 30 %. Aside from VS5/ULOQ, the intra-assay accuracy across the six runs ranged from -17.4 to 4.2 % bias; inter-assay accuracy varied from -10.9 to -1.8 % bias.

### Selectivity

Ten individual AD CSF samples were tested for recovery of a known spike concentration of synthetic calibrator. A spike concentration approximately equal to VS3 was selected to ensure that all tested samples would remain in range despite anticipated variation in endogenous A $\beta$ <sub>42</sub> concentration. Acceptance criteria for this experiment was  $\leq\pm 20$  % bias. The endogenous analyte levels ranged from 483 to 1,950 pg/mL. Across that range, the spike was recovered to an acceptable level (86.1 to 98.8 %) regardless of endogenous A $\beta$ <sub>42</sub> concentration (see Supplementary Table I).

### Specificity

The ability of the assay to distinguish between A $\beta$ <sub>42</sub> and A $\beta$ <sub>40</sub> was tested in both VS2 (designated LQC after the core runs) and VS4 (HQC). Both QC samples were tested in the presence of 0, 900, 2,250, and 4,500 pg/mL A $\beta$ <sub>40</sub>. The acceptance criterion for this experiment was set to  $\leq\pm 20$  % bias from established A $\beta$ <sub>42</sub> concentration. In all cases, A $\beta$ <sub>42</sub> recovery was acceptable (range 86.1–98.8 %), with no correlation between recovery and spike concentration (see Supplementary Fig. 5).

### Parallelism

Three individual AD CSF samples were diluted from neat (in this experiment, MRD was not performed) to 1 in 8 to assess the parallelism between the endogenous analyte in CSF and the synthetic calibrators in assay buffer. As in method development, the A $\beta$ <sub>42</sub> reported concentration is suppressed in 100 and 50 % CSF (see Fig. 5). In this

**Table I.** Precision and Accuracy of the Method

VS level	Intra-assay precision (%CV)	Inter-assay precision (%CV)	Intra-assay accuracy (%bias)	Inter-assay accuracy (%bias)
VS1 (LLOQ)	1.5 to 4.9	7.0	-12.8 to 4.2	-7.9
VS2 (LQC)	3.1 to 5.5	6.1	-16.7 to -4.0	-8.3
VS3 (MQC)	1.7 to 3.6	4.2	-6.0 to 2.9	-1.8
VS4 (HQC)	0.6 to 3.1	6.6	-17.4 to -1.4	-10.9
VS5 (ULOQ)	1.0 to 8.4	12.1	-22.1 to 12.9	-6.3

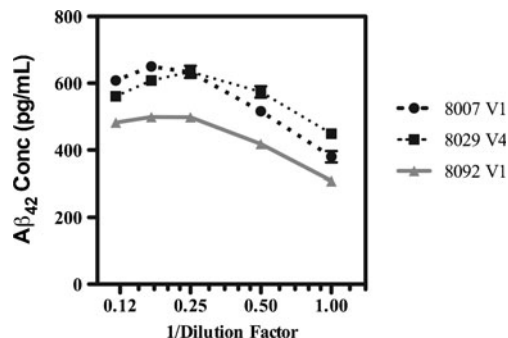
experiment, dilutions of 1 in 4, 1 in 6, and 1 in 8 all generate a similar back-calculated concentration, indicating that the assay could run at an MRD of four or six, and that the assay demonstrates parallelism at CSF concentrations of 25 % or below.

### Linearity of Dilution

To test for dilutional linearity and the presence of a hook effect, LQC was spiked to 9,000 pg/mL with A $\beta$ <sub>42</sub> calibrator and then diluted in several increments up to 100-fold. All samples with a nominal A $\beta$ <sub>42</sub> concentration  $\geq$ 4,500 pg/mL generated a signal above the ULOQ, indicating that no hook effect occurs at high A $\beta$ <sub>42</sub> concentration. The dilutions with nominal A $\beta$ <sub>42</sub> concentration values within the quantitative range of the assay displayed a linear relationship between expected and observed A $\beta$ <sub>42</sub> concentration (see Supplementary Fig. 6). These results indicate that the assay displays acceptable dilutional linearity within the quantitative range and no hook effect.

### Interference from Hemolysis

The measured A $\beta$ <sub>42</sub> concentration in two lots of AD donor CSF spiked with three levels of lysed blood was within 20 % of the appropriate mock-treated control (data not shown). Therefore, the method was deemed free from interference from hemolyzed blood up to the highest contamination level (10 %) tested.



**Fig. 5.** Parallelism of calibrator and CSF sample concentration response. Three lots of individual AD donor CSF were diluted in assay buffer to various levels prior to loading into the assay plate. The back-calculated A $\beta$ <sub>42</sub> concentration for each lot at each dilution level was plotted against the applied dilution factor. Error bars describe the range around the mean of the two replicates

<sup>2</sup> Due to the relative nature of the accuracy determinations (see footnote 1), the total error calculations are also not absolute. The term is used to assess the cumulative variance of the method, considering both the absolute precision value and the relative accuracy value, for comparison to other published biomarker validations in which the same relative accuracy qualification generally applies (see “Discussion”).

### Stability

LQC and HQC in LoBind tubes and two individual ISs in Nunc cryovials were stability tested at room temperature, on ice, after several freeze–thaw cycles, and at  $-70^{\circ}\text{C}$ . Benchtop stability at room temperature was established for both QCs in LoBind tubes and ISs in Nunc cryovials to 21 h in the absence of Tween-20 and 5 h in the presence of Tween-20. Benchtop stability on ice was established for both QCs in LoBind tubes and ISs in Nunc cryovials to 22.5 h in the absence of Tween-20 and 4 h in the presence of Tween-20. Incurred samples in Nunc vials were tolerant to three cycles of freeze–thaw (additional cycles were not tested). Long-term frozen stability studies in LoBind and Nunc has been established to 3 months at  $-70^{\circ}\text{C}$  (studies ongoing). These studies indicate that  $\text{A}\beta_{42}$  is sufficiently stable in CSF over timeframes appropriate for assay execution, with sufficient tolerance for unexpected variations from defined procedure.

## DISCUSSION

### General Observations on the Use of $\text{A}\beta_{42}$ Assays in the Field

The Innogenetics INNOTEST®  $\text{A}\beta_{42}$  ELISA kit has been commercially available since 2000 and is one of the most widely used kits for CSF  $\text{A}\beta_{42}$  measurement. It has been successfully used to categorize subjects by disease status (3,4,15,16) and has the potential to be an FDA-approved diagnostic test for AD. There is also growing interest in using these types of kits to measure CSF  $\text{A}\beta_{42}$  changes as a pharmacodynamic biomarker of drug effect in clinical trials (23,24). Since such pharmacodynamic changes may be subtle, excellent assay precision and reproducibility are highly desirable.

In order to support our own clinical efforts, we have, to our knowledge, performed the first GLP-equivalent advanced validation of this (or any other  $\text{A}\beta_{42}$ ) assay. During pre-validation method development, we systematically examined every aspect of the method, and made significant modifications when required to fulfill advanced validation requirements. A summary of the original manufacturer-recommended assay parameters, and the various modifications we made to support advanced validation can be found in Table II. These modifications, and the advanced assay validation that followed, are discussed in more detail below.

### Establishment of Minimum Required Dilution for CSF in INNOTEST ELISA

The INNOTEST® manufacturer recommends that CSF samples be analyzed neat, and the field has followed this recommendation in establishing apparent  $\text{A}\beta_{42}$  levels in AD, MCI and cognitively normal subjects (e.g., 3,9,15,16,25). However, we have demonstrated in this report that dilution of CSF results in increased  $\text{A}\beta_{42}$  measured concentrations; such apparent signal suppression in neat CSF was also demonstrated in this assay recently by Bjerke *et al.* (17), although an MRD was not determined. Due to the adequate sensitivity of the method we opted to consider MRD as the dilution of matrix that fully mitigates the impact of matrix on the back-calculated analyte value (in this case, a dilution factor of 6). Lower dilution factors

can be considered if concerns about sensitivity outweigh concerns about complete recovery of  $\text{A}\beta_{42}$  in the CSF (however, the precision of the assay would need to be re-established). To our knowledge, the present study is the first to rigorously assign an MRD for this assay in CSF.

### Establishment of Accurate Reference Range for $\text{A}\beta_{42}$ in AD CSF

A critical implication of our MRD analysis is that published studies using neat samples to establish reference ranges in normal and diseased human populations likely underestimate true  $\text{A}\beta_{42}$  concentrations. These studies (for example, 3–5,15,16) utilizing the INNOTEST®  $\text{A}\beta_{42}$  ELISA reported that human  $\text{A}\beta_{42}$  CSF concentrations average approximately 560–760 pg/mL in normal subjects and approximately 380–490 pg/mL in AD. In order to use the most appropriate calibration curve for our study, we analyzed multiple incurred mild AD study samples after MRD/Tween treatment and observed an  $\text{A}\beta_{42}$  range of 1,500–3,500 pg/mL. Further investigation of multiple normal and AD CSF samples comparing this modified INNOTEST® assay with other methodologies (once validated to a similarly advanced level) would provide a greater understanding of the true  $\text{A}\beta_{42}$  concentration range in humans and would provide an indication of the absolute accuracy of each validated method.

### Modification of Calibration Curve to Correctly Fit Accurate Reference Range

The establishment of an MRD in this method effectively expanded the quantitative range of the method 6-fold, as diluted CSF samples were queried against neat, buffer-based calibrators. In order to optimally orient calibrators and QC samples in relation to the established reference range, the calibration range was adjusted as defined in Table II. The quantitative range of the optimized method was designed to accommodate a substantial increase or decrease in CSF  $\text{A}\beta_{42}$  levels due to a therapeutic effect during a clinical trial (see Fig. 3). This modified calibration series performed well in validation and is recommended for future users of the method, coupled with an MRD of 6.

### Detergent Mitigation of $\text{A}\beta_{42}$ Adsorption

Studies performed by Innogenetics and others (20) (Vanderstichele *et al.*, personal communications) indicate that  $\text{A}\beta_{42}$  can adsorb significantly and variably to polypropylene tubes made by different manufacturers. To respond to loss of  $\text{A}\beta_{42}$  in our clinical sample storage vials, we used a method introduced by McCush *et al.* (22) in which addition of a final concentration of 0.2 % Tween-20 can recover adsorbed peptide. Furthermore, we were able to extend those findings by showing complete mitigation of  $\text{A}\beta_{42}$  adsorption, using non-adsorbing LoBind tubes as comparators. The chosen level of Tween-20 does not interfere with assay performance (Supplementary Fig. 4), and therefore this post-collection mitigation strategy is suggested for use across a wide range of polypropylene tubes and for existing samples already in long-term storage.

**Table II.** Method Development Aspects and Improvements to the Assay

Parameter	Per INNOTEST A $\beta$ <sub>1-42</sub> kit protocol	Modified A $\beta$ <sub>42</sub> method	Advantage of modification
Calibration range	2,000–125 pg/mL	1,000–40 pg/mL	Better fit to reference range
Number of calibrators	6	10	Improved definition of 4-PL regression
Non-linear regression method	Sigmoidal curve fit recommended	4-PL with 1/Y weighting	Increased accuracy at low analyte concentration
Quantitative range	Undefined	375–4,500 pg/mL	Provides a defined range with which to plan clinical sample analysis
MRD	1 (analyze sample neat)	6	Increases accuracy of quantitation
Sensitivity	LOD is $\pm$ 50 pg/mL	LLOQ=62.5 pg/mL (after MRD, 375 pg/mL)	Established sensitivity defines assay utility in particular clinical populations
VS/QCs	None	5 VS levels/3 QC levels (composed of AD donor CSF)	Provides ability to monitor assay performance and establish validation parameters
Sample controls	None	2 individual AD donor CSF samples	Ensures the robustness of the assay with unadulterated CSF
Sample pre-treatment	None	Spike sample with Tween-20 to 0.2 % final	Mitigates potential adsorption of analyte to sample collection tube
Assay acceptance criteria	None	4–6–20 acceptance criterion on QCs	Establishes mechanism for the unbiased rejection of poorly performing plates

### Appropriate Source of VS/QC Material

In their description of fit-for-purpose biomarker development and validation, Lee *et al.* (21) state that VS and QC material “should be as closely related to the study samples as possible.” We followed these recommendations by sourcing enough CSF from human AD subjects to create large pools for both assay validation and sample analysis. This CSF was collected, processed and stored using defined and standardized procedures and therefore was deemed superior to widely available remnant CSF. In addition, AD CSF contains lower A $\beta$ <sub>42</sub> concentrations, enabling the creation of VS/QC pools from purely native CSF to optimally control the performance of the method during clinical sample analysis. However, to establish the full quantitative range of the method and to allow for potential modulation of A $\beta$ <sub>42</sub> concentration upon therapy, adulteration of the CSF was necessary: HQC and ULOQ were created by spiking in synthetic A $\beta$ <sub>42</sub> and LLOQ was generated by dilution of CSF in assay buffer. Both manipulations (spiking or dilution) were supported by the MRD/Parallelism experiments (Figs. 1 and 5), in that synthetic analyte possessed a similar concentration/signal correlation to endogenous A $\beta$ <sub>42</sub> and back-calculated sample concentrations remain constant at CSF levels below 16.7 %. We believe that the use of highly relevant human CSF for method validation and QC, represents a significant improvement over buffer-based or remnant CSF approaches.

### Performance of the Optimized Method

The process of method validation revealed that the optimized method is suitable for the detection of biomarker shifts during a clinical study. The intra-assay precision throughout the quantitative range was determined to be below 10 % and both the intra- and inter-assay precision results compare well to commonly employed acceptance criteria for biomarker assays (26). As discussed above, accuracy could be validated only to a relative level due to the presence of endogenous analyte in all human CSF samples. Although absolute accuracy is ultimately

not required for use of the method to assess therapeutic effects, the mitigation of two major sources of inaccuracy as reported here represents significant progress toward that goal. Further development of the method as a diagnostic tool would require cross validation with other validated quantitative techniques in order to confirm absolute accuracy. Despite the limitations imposed by the matrix in question, the accuracy results and total error calculations generated during method validation revealed the consistency of quantitation across the entire quantitative range and compare well to the biomarker field (26). The optimized method is selective and specific for A $\beta$ <sub>42</sub>, is free of the artifacts of non-parallelism, dilutional non-linearity, and hook effect, and is tolerant of significant blood contamination of CSF. The CSF-based QC samples introduced into the method are stable under conditions expected during normal assay execution, and CSF samples can be stored at  $-70^{\circ}\text{C}$  for at least 3 months.

### Implications for the AD Field

In addition to diagnosing and stratifying patients, there is an increasing interest in using CSF A $\beta$ <sub>42</sub> changes as a potential marker of drug efficacy, therefore requiring a reliable, accurate and precise A $\beta$ <sub>42</sub> assay. However, the lack of rigorously validated methods, as well as the documented variability both between and within commonly used A $\beta$ <sub>42</sub> assay platforms has impeded the use of these assays in drug trials. In this present study, we have demonstrated that the INNOTEST A $\beta$ <sub>42</sub> assay meets advanced validation criteria when an MRD is incorporated, and is thus useful for measuring drug induced biomarker changes. However, the different absolute values between platforms create uncertainty in the absolute accuracy of each method. In several recent reports, QC human CSF samples were tested in three commonly used assay kits: INNO-BIA AlzBio3® (a Luminex bead-based method), the MSD® MULTI-SPOT® Human (6E10) Abeta Triplex Assay, and the INNOTEST A $\beta$ <sub>42</sub> ELISA (17,27,28). Marked differences in A $\beta$ <sub>42</sub> concentration within the same CSF sample were revealed when the methods were compared, with some differences of greater than 3-fold reported. The increase in



A $\beta_{42}$  concentration upon dilution to MRD and adsorption mitigation revealed in this report raises the distinct possibility that matrix effects and/or adsorbance contribute to the observed inter-platform discordance. Careful validation of all CSF A $\beta_{42}$  assays would surely advance the effort to harmonize the clinical application of these technologies.

## CONCLUSIONS

In conclusion, we have conducted the first GLP-level method development and advanced validation of a method for the quantitation of A $\beta_{42}$  in human CSF. Mitigation of both matrix interference and analyte adsorption were critical to a successful validation. In addition, we have conducted this validation using CSF from individual AD patients for all critical experiments including precision and accuracy. This modified ELISA assay can be used with confidence to precisely and accurately measure CSF A $\beta_{42}$  in current and future clinical studies of novel therapeutic modalities.

## ACKNOWLEDGMENTS

We thank Mike Bleavins, John Allinson, and Andy Jones for helpful discussions.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## REFERENCES

- Apostolova LG, Hwang KS, Andrawis JP, Green AE, Babakchani S, Morra JH, *et al.* 3D PIB and CSF biomarker associations with hippocampal atrophy in ADNI subjects. *Neurobiol Aging*. 2010;31:1284–303.
- Bouwman FH, Schoonenboom NS, Verwey NA, van Elk EJ, Kok A, Blankenstein MA, *et al.* CSF biomarker levels in early and late onset Alzheimer's disease. *Neurobiol Aging*. 2009;30:1895–901.
- Engelborghs S, Maertens K, Vloeberghs E, Aerts T, Somers N, Marien P, *et al.* Neuropsychological and behavioural correlates of CSF biomarkers in dementia. *Neurochem Int*. 2006;48:286–95.
- Maruyama M, Arai H, Sugita M, Tanji H, Higuchi M, Okamura N, *et al.* Cerebrospinal fluid amyloid beta(1–42) levels in the mild cognitive impairment stage of Alzheimer's disease. *Exp Neurol*. 2001;172:433–6.
- Parnetti L, Lanari A, Silvestrelli G, Saggese E, Reboldi P. Diagnosing prodromal Alzheimer's disease: role of CSF biochemical markers. *Mech Ageing Dev*. 2006;127:129–32.
- Shaw LM, Vanderstichele H, Knapiak-Czajka M, Clark CM, Aisen PS, Petersen RC, *et al.* Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects. *Ann Neurol*. 2009;65:403–13.
- Diniz BS, Pinto Junior JA, Forlenza OV. Do CSF total tau, phosphorylated tau, and beta-amyloid 42 help to predict progression of mild cognitive impairment to Alzheimer's disease? A systematic review and meta-analysis of the literature. *World J Biol Psychiatry*. 2008;9:172–82.
- Fjell AM, Walhovd KB, Fennema-Notestine C, McEvoy LK, Hagler DJ, Holland D, *et al.* CSF biomarkers in prediction of cerebral and clinical change in mild cognitive impairment and Alzheimer's disease. *J Neurosci*. 2010;30:2088–101.
- Hampel H, Teipel SJ, Fuchsberger T, Andreasen N, Wiltfang J, Otto M, *et al.* Value of CSF beta-amyloid1–42 and tau as predictors of Alzheimer's disease in patients with mild cognitive impairment. *Mol Psychiatry*. 2004;9:705–10.
- Ivanioiu A, Sindic CJ. Cerebrospinal fluid TAU protein and amyloid beta42 in mild cognitive impairment: prediction of progression to Alzheimer's disease and correlation with the neuropsychological examination. *Neurocase*. 2005;11:32–9.
- Mattsson N, Zetterberg H, Hansson O, Andreasen N, Parnetti L, Jonsson M, *et al.* CSF biomarkers and incipient Alzheimer disease in patients with mild cognitive impairment. *JAMA*. 2009;302:385–93.
- Visser PJ, Verhey F, Knol DL, Scheltens P, Wahlund LO, Freund-Levi Y, *et al.* Prevalence and prognostic value of CSF markers of Alzheimer's disease pathology in patients with subjective cognitive impairment or mild cognitive impairment in the DESCRIPA study: a prospective cohort study. *Lancet Neurol*. 2009;8:619–27.
- Hampel H, Buerger K, Zinkowski R, Teipel SJ, Goernitz A, Andreasen N, *et al.* Measurement of phosphorylated tau epitopes in the differential diagnosis of Alzheimer disease: a comparative cerebrospinal fluid study. *Arch Gen Psychiatry*. 2004;61:95–102.
- Hampel H, Ewers M, Burger K, Annas P, Mortberg A, Bogstedt A, *et al.* Lithium trial in Alzheimer's disease: a randomized, single-blind, placebo-controlled, multicenter 10-week study. *J Clin Psychiatry*. 2009;70:922–31.
- Bouwman FH, van der Flier WM, Schoonenboom NS, van Elk EJ, Kok A, Rijmen F, *et al.* Longitudinal changes of CSF biomarkers in memory clinic patients. *Neurology*. 2007;69:1006–11.
- Maddalena A, Papassotiropoulos A, Muller-Tillmanns B, Jung HH, Hegi T, Nitsch RM, *et al.* Biochemical diagnosis of Alzheimer disease by measuring the cerebrospinal fluid ratio of phosphorylated tau protein to beta-amyloid peptide42. *Arch Neurol*. 2003;60:1202–6.
- Bjerke M, Portelius E, Minthon L, Wallin A, Anckarsater H, Anckarsater R, *et al.* Confounding factors influencing amyloid beta concentration in cerebrospinal fluid. *Int J Alzheimers Dis*. 2010;2010:986310.
- Lewczuk P, Beck G, Esselmann H, Bruckmoser R, Zimmermann R, Fiszer M, *et al.* Effect of sample collection tubes on cerebrospinal fluid concentrations of tau proteins and amyloid  $\beta$  peptides. *Clin Chem*. 2006;52:332–4.
- Andreasen N, Hesse C, Davidsson P, Minthon L, Wallin A, Winblad B, *et al.* Cerebrospinal fluid  $\beta$ -amyloid<sub>(1–42)</sub> in Alzheimer disease. *Arch Neurol*. 1999;56:673–80.
- Pica-Mendez AM, Tanen M, Dallob A, Tanaka W, Laterza OF. Nonspecific binding of A $\beta_{42}$  to polypropylene tubes and the effect of Tween-20. *Clin Chim Acta*. 2010;411:1833.
- Lee JW, Devanarayan V, Barrett YC, Weiner R, Allinson J, Fountain S, *et al.* Fit-for-purpose method development and validation for successful biomarker measurement. *Pharm Res*. 2006;23:312–28.
- McCush F, Raha N, Durham R, Chappell J, Sharma V, Tew Z, Lane D. Validation of Alzheimer biomarker assays using various analytical techniques in CSF and plasma. In 9th Annual Marian S Ware Neurodegeneration Research Retreat (Current Research in Alzheimer Biomarkers; 2009 October 16, 2009).
- Siemers E, DeMattos RB, May PC, Dean RA. Role of biochemical Alzheimer's disease biomarkers as end points in clinical trials. *Biomark Med*. 2010;4:81–9.
- Zetterberg H, Mattsson N, Shaw LM, Blennow K. Biochemical markers in Alzheimer's disease clinical trials. *Biomark Med*. 2010;4:91–8.
- Parnetti L, Lanari A, Saggese E, Spaccatini C, Gallai V. Cerebrospinal fluid biochemical markers in early detection and in differential diagnosis of dementia disorders in routine clinical practice. *Neurol Sci*. 2003;24:199–200.
- Viswanathan CT, Bansal S, Booth B, DeStefano AJ, Rose MJ, Sailstad J, *et al.* Quantitative bioanalytical methods validation and implementation: best practices for chromatographic and ligand binding assays. *Pharm Res*. 2007;24:1962–73.
- Fagan AM, Shaw LM, Xiong C, Vanderstichele H, Mintun MA, Trojanowski JQ, *et al.* Comparison of analytical platforms for cerebrospinal fluid measures of [beta]-Amyloid 1–42, total tau, and P-tau181 for identifying Alzheimer disease amyloid plaque pathology. *Arch Neurol*. 2011;68:1137–44.
- Mattsson N, Andreasson U, Persson S, Arai H, Batish SD, Bernardini S, *et al.* The Alzheimer's association external quality control program for cerebrospinal fluid biomarkers. *Alzheimers Dement*. 2011;7:386.e6–95.e6.