Clin Chem Lab Med 2006;44(12):1472–1480 © 2006 by Walter de Gruyter • Berlin • New York. DOI 10.1515/CCLM.2006.258

Analytical performance and clinical utility of the INNOTEST® PHOSPHO-TAU$_{(181P)}$ assay for discrimination between Alzheimer's disease and dementia with Lewy bodies

Hugo Vanderstichele[1,*], Karen De Vreese[1,a], Kaj Blennow[2], Niels Andreasen[3,a], Christian Sindic[4,a], Adrian Ivanoiu[4,a], Harald Hampel[5,a], Katharina Bürger[5,a], Lucilla Parnetti[6,a], Alessia Lanari[6,a], Allesandro Padovani[7,a], Monica DiLuca[7,a], Miriam Bläser[1,a], Annika Ohrfelt Olsson[2], Hans Pottel[1], Frank Hulstaert[1,a] and Eugeen Vanmechelen[1]

[1] Innogenetics NV, Gent, Belgium
[2] Sahlgren's University Hospital, Mölndal, Sweden
[3] Neurotec, Department of Geriatric Medicine, Karolinska University Hospital, Stockholm, Sweden
[4] Cliniques Universitaires Saint-Luc, Brussels, Belgium
[5] Ludwig Maximilian University, Department of Psychiatry, Munich, Germany
[6] University of Perugia, Perugia, Italy
[7] Università degli Studi Brescia, Brescia, Italy

Abstract

Background: Total tau (T-tau) and β-amyloid$_{(1-42)}$ (Aβ$_{1-42}$) levels in cerebrospinal fluid (CSF) can differentiate Alzheimer's disease (AD) from normal aging or depressive pseudo-dementia. Differential diagnosis from dementia with Lewy bodies (DLB) in clinical settings is difficult.
Methods: The analytical performance of the INNOTEST® PHOSPHO-TAU$_{(181P)}$ assay was validated in terms of selectivity, sensitivity, specificity, precision, robustness, and stability. Clinical utility of the assay alone, or combined with T-tau and Aβ$_{1-42}$, for discrimination of AD (n=94) from patients suffering from DLB (n=60) or from age-matched control subjects (CS) (n=60) was assessed in a multicenter study.
Results: CSF concentrations of tau phosphorylated at threonine 181 (P-tau$_{181P}$) in AD was significantly higher than in DLB and CS. Discriminant analysis, a classification tree, and logistic regression showed that P-tau$_{181P}$ was the most statistically significant single variable of the three biomarkers for discrimination between AD and DLB.
Conclusions: P-tau$_{181P}$ quantification is a robust and reliable assay that may be useful in discriminating AD from DLB.
Clin Chem Lab Med 2006;44:1472–80.

Keywords: Alzheimer's disease; cerebrospinal fluid; dementia; diagnosis; performance; P-tau$_{181P}$.

[a]These authors contributed to only the clinical part of the study.

*Corresponding author: Hugo Vanderstichele, PhD, Innogenetics, Industriepark Zwijnaarde 7, P.O. Box 4, 9052 Gent, Belgium
Phone: +32-9-2410711, Fax: +32-9-2410907,
E-mail: hugovdr@innogenetics.be

Introduction

Alzheimer's disease (AD) can be considered the most important neurodegenerative brain disease because of its frequent occurrence and devastating consequences, affecting up to 10% of the population older than 65 years. Its impact is increasing in parallel with the general worldwide aging of the population. Histologically, brains of AD patients are characterized by the presence of neurofibrillary tangles (NFT), senile plaques, and synapse loss (1). The major component of the NFT structures is hyperphosphorylated tau protein. Tau protein is a normal brain phosphoprotein that binds to microtubules in neuronal axons, thereby promoting neural assembly and stability. When hyperphosphorylated, tau loses its stabilizing function, resulting in axon instability and reduced intracellular transport. More than 70 potential phosphorylation sites have been identified in the tau molecule (2). Tau and phospho-tau are closely associated with the pathology of several age-related neurodegenerative disorders, such as AD, frontotemporal lobe dementia (FTD), and dementia with Lewy bodies (DLB). More reliable diagnostic procedures are required, since AD is only correctly diagnosed in clinical settings in 65%–90% of cases (3). The higher percentage is especially associated with academic settings with clinical follow-up, while the lower figure more likely occurs in non-academic settings and at first clinical work-up. Consensus criteria for AD biomarkers have been published (4).

Earlier studies indicated that tau phosphorylated at threonine 181(P-tau$_{181P}$) in cerebrospinal fluid (CSF) might be used to discriminate AD from DLB (5). Since distinction of AD from DLB is difficult in clinical settings, but important for patient management, as the treatment strategies are different, P-tau$_{181P}$ determination may prove to be of clinical value. In addition, combined measurement of β-amyloid$_{(1-42)}$ (Aβ$_{1-42}$) and P-tau$_{181P}$ levels has been shown to differentiate AD from FTD (6).

We describe the performance characteristics of the INNOTEST® PHOSPHO-TAU$_{(181P)}$ assay and an evaluation of its clinical utility, alone or combined with total tau (T-tau) and Aβ$_{1-42}$ for the discrimination of AD from DLB and control subjects (CS).

Materials and methods

Assays

The INNOTEST® PHOSPHO-TAU$_{(181P)}$ assay is a solid-phase enzyme immunoassay in which P-tau protein is captured by monoclonal antibody (mAb) HT7 (Innogenetics, Gent, Belgium). A sample of 75 μL of undiluted CSF is added and

incubated overnight at 2–8°C with 25 μL of the biotinylated anti-P-tau$_{181P}$ mAb AT270 (Innogenetics). Any antigen-antibody complex formed is then detected by peroxidase-labeled streptavidin (SV-PO). Two wash steps are included in the test instructions: one after incubation of the antigen and biotinylated detector antibody, and the other after SV-PO. No difference was observed between a manual wash procedure or an automated wash procedure on a Columbus instrument (data not shown). The assay contains a sample addition monitoring system to confirm the addition of CSF or SV-PO. A phosphorylated peptide, containing both the epitope of HT7 (epitope P_{159}PGQK$_{163}$, numbered according to the longest tau isoform) (7) and AT270 (epitope P_{176}PAPKTP$_{182}$), was used for calibration (peptide sequence, acetyl-P_{154}RGAAPPGQKGQANATRIPAKTPPAPK**T**PPSSGE$_{187}$; molecular weight, 3455 Da; obtained from Neosystems, Strasbourg, France). The calibrator peptide should be considered as a reference standard. Its structure and size differ from the size of the native P-tau protein. The assay format quantifies all tau isoforms phosphorylated at threonine 181. Full details of the assays used in the present study have been described previously (8–10).

Production process

Plate coating The process of mAb coating of immunoplates (Nunc Maxisorps™, San Diego, CA, USA) was fully validated. The homogeneity of the mAb coating was verified for each production run by performing an assay with 125 pg/mL of the calibrator peptide. For this purpose, immunoplates were assembled containing 8-well strips obtained from plates selected over the whole production run. Two combined plates were used for testing. Half of the plate (either the upper or lower part) was filled with 125 pg/mL of the calibrator (positive sample) or the blank (negative sample). The percentage coefficient of variation (%CV) for the 96 positive samples was calculated. The effect of the blocking and fixation stabilization process on P-tau$_{181P}$ concentrations measured in CSF was verified for 38 CSF samples (concentration range 3.5–207.4 pg/mL; calculated on fresh plates) by comparison of results obtained on fresh and stabilized immunoplates, as assessed by a paired two-tailed t-test on log-transformed values.

Stability The stability of the whole kit was evaluated on kits from two production runs. All components of the kit were stored at 2–8°C for more than 24 months. Calibrator peptides were stored at 2–8°C or −20°C. At each time point, assays were performed according to the test instructions. Three control samples (run validation samples) were prepared by addition of calibrator peptides to artificial CSF (ACSF) (composition NaCl, 138 mM; KCl, 2.8 mM; $CaCl_2 \cdot 2H_2O$, 1.05 mM; $MgCl_2 \cdot 6H_2O$, 1.15 mM; $NaH_2PO_4 \cdot H_2O$, 1.25 mM; glucose, 0.06% w/w; and human serum albumin, 0.16% w/w).

The effect of freezing (up to ten times) on the prepared calibrator concentration range was also evaluated. At each time point, three aliquots of each calibrator concentration were taken and analyzed on two different mAb-coated plates. Each vial was stored for at least 1 day after each freezing cycle at −20°C. All samples were analyzed in the same experiment.

Analytical performance

Curve fitting Experiments were performed to select the most appropriate mathematical model for curve fitting, as well as the number and concentration of dilutions. A lack-of-fit test determined whether the predicted sigmoidal calibration curve adequately described the data obtained. Evaluation was performed based on model fit (R^2 values), residual analysis, and analysis of the variability on inverse prediction (e.g., estimating concentrations for unknown samples).

Selectivity and specificity Selectivity (ability of the assay to measure and differentiate P-tau$_{181P}$ in the presence of matrix components that may be expected to be present) and specificity (detection of P-tau$_{181P}$ as opposed to other P-tau proteins) were evaluated using synthetic biotinylated peptides (sequence Q_{165}ANATRIPAKTPPAPK**T**PPSSGEPPKS$_{191}$) either phosphorylated at Thr$_{181}$ or non-phosphorylated. For this purpose, SV-coated immunoplates were first incubated for 1 h with biotinylated peptides. After a wash step, mAbs were incubated overnight at 2–8°C. Detection of the bound mAbs was achieved using peroxidase-labeled goat anti-mouse antibodies (Jackson Immunoresearch Laboratories, Lucron Bioproducts; De Pinte, Belgium).

Analytical sensitivity Blank samples (bl) (~sample diluent) were used to calculate the limit of detection (LOD; $mean_{bl} + 2\ SD_{bl}$) and lower limit of quantitation (LLOQ; $mean_{bl} + 10\ SD_{bl}$). The upper limit of quantitation (ULOQ) was determined from linearity and accuracy results.

Accuracy Using data from the precision experiments, the percentage bias (known value−predicted value/known value×100%) was calculated. However, in principle, accuracy (closeness of test results obtained by the method to the true concentration of the analyte) was difficult to define in this case, since no reference material (gold standard) is currently available.

Repeatability and reproducibility Intermediate precision (closeness of individual measurements of an analyte when the procedure is applied repeatedly) was evaluated by the analysis of 12 samples (including six points for calibration and two blanks) for eight replicates on three coated ELISA plates from five different production runs. For each plate and sample, the average, SD, CV, and bias were calculated using the calibrator curve provided in the kit.

Repeatability and reproducibility were tested according to ISO 5725-2 guidelines (11). Control samples were analyzed by three different operators in quadruplicate, using three different vials of calibrators on ELISA plates from the same production run (repeatability), or by one operator in quadruplicate using the same calibrators on plates from three different production runs (reproducibility). The method performance chart presents inaccuracy (percentage bias) plotted against the imprecision (CV).

Linearity Assay linearity (ability of the method to elicit test results that are directly proportional to the analyte concentration) was evaluated by combining samples with high or low concentrations of P-tau$_{181P}$ in different volume ratios (1:0, 0.75:0.25, 0.5:0.5, 0.25:0.75 and 0:1). Results calculated for the mixed samples were plotted against predicted concentrations obtained in undiluted CSF.

Test robustness The robustness (capacity of the assay to remain unaffected by deliberate variations in method parameters) or effect of variation in test parameters (antigen and detector-antibody incubation time; SV-PO incubation time, SV-PO incubation temperature, substrate incubation time, and substrate incubation temperature) on assay performance was characterized by factorial design of experiments. In a full factorial design, all combinations of boundary assay parameter settings were tested on ACSF

Table 1 Assay parameters tested.

	Boundary settings		
	Low	Standard	High
Antigen incubation time, h	14	16	18
SV-PO incubation time, min	55	60	65
SV-PO incubation temperature, °C	18	24	30
Substrate incubation time, min[a]	27	30	33
Substrate incubation temperature, °C	18	24	30

[a] For substrate incubation, only the combined low and high settings were included in combinations with settings of the other factors.

samples, spiked with different concentrations of P-tau$_{181P}$-peptide (Table 1). Each sample was tested in quadruplicate.

CSF sample preparation To assess the effect of centrifugation, CSF samples (n = 15, concentration range 20–160 pg/mL P-tau$_{181P}$) were either aliquoted immediately after collection and stored at −80°C, or centrifuged at 2000 *g* for 10 min at 4°C and subsequently frozen at −80°C. A Wilcoxon matched-pairs test was used to compare both methods.

To assess the effect of storage conditions, CSF samples (n = 11) were aliquoted and stored immediately at −80°C, or stored for 4, 24, or 72 h at 4°C or at room temperature before freezing at −80°C. Repeated-measures ANOVA was used to evaluate the effects of temperature storage conditions.

To evaluate the impact of freezing, CSF samples (n = 8) stored at −80°C were thawed and refrozen several times. Each CSF sample was stored at −80°C for at least 1 night. The effect of additional freezing of CSF was evaluated by repeated-measures ANOVA.

Clinical utility using stored patient samples

Samples The multicenter study used previously stored CSF samples available for research purposes from five laboratories. No follow-up samples were included in the study. Patients suffering from relevant concomitant diseases, such as brain tumors or metastases, CSF infection or stroke, were excluded. The AD group (n = 94) consisted of randomly selected patients diagnosed with probable AD according to DSM-IV criteria and criteria established by the work group of the National Institute of Neurological and Communicative Disorders and Stroke/Alzheimer's Disease and Related Disorders Association (NINCDS-ADRDA) (12). DLB patients (n = 60) were randomly selected and diagnosed according to the criteria of McKeith et al. (13). Mini-mental state examination (MMSE) (14) scores were documented for most patients and for some of the CS. Clinical diagnosis preceded or coincided with CSF collection. CS (n = 60) were defined as either (i) apparently healthy (n = 37) or (ii) subjects without systemic inflammatory or central nervous system disorders, aged over 55 years, who had undergone lumbar puncture and CSF collection for diagnosis of mechanical problems (n = 6; 2 nucleopathies, 1 headache in epilepsy, 1 disk herniation, 1 cervical lumbar arthrosis, 1 minor head trauma) or (iii) depression (n = 17). Only CSF samples collected in the 2 years preceding the study start, stored in polypropylene tubes, and kept (pending biochemical analysis) frozen at −20°C or lower were included. The number of freeze/thaw cycles was recorded.

All clinical samples were tested in duplicate at the respective centers (multicenter collection and testing). Operators were blinded to the expected test outcome in terms of clinical diagnosis. Concentrations of P-tau$_{181P}$, T-tau, and Aβ_{1-42} were required for all samples. When historical results of INNOTEST® hTAU Ag and INNOTEST® β-AMYLOID$_{(1-42)}$ assays were available, the investigator was free to decide whether or not to repeat these analyses. As such, results for T-tau and Aβ_{1-42} need to be interpreted with some caution, as they were partly based on historical results not obtained in the standardized protocol setting.

The study presented no risk(s) or harm to the subjects. Previously stored samples available for research purposes were used for this study. Ethics Committee approval to conduct the study was not required for three centers, was obtained in one center for research purposes, and was given for the fifth center. Subject confidentiality was maintained at all times, within the legal context of the country, by removal of all direct subject identifiers from patient samples and clinical documentation submitted for this study. Anonymity was also assured in any report or publication.

Statistics

Several statistical techniques [receiver operating characteristic curve (ROC) analysis, discriminant analysis, classification tree, and logistic regression] were used in an attempt to optimally classify AD, DLB, and CS in the correct group based on biomarker information. Transformed variables were employed if this resulted in increased discriminative power. A Wilcoxon rank sum test was used to compare group differences. Spearman's correlation was used to evaluate correlation of P-tau$_{181P}$ and age or MMSE. The data for each assay were analyzed using sigmoidal curve fitting with the GraphPad Prism program (Version 4.01; GraphPad Software, San Diego, CA, USA).

Results

Production process

The homogeneity of the coating process was analyzed by performing an assay with 125 pg/mL of the calibrator peptide on immunoplates obtained from ten different production runs. The median CV (p10–p90) for the positive sample was 3.8% (2.6%–5.8%). The process of blocking and fixation of mAbs on the plates did not affect the P-tau$_{181P}$ concentration in CSF, as verified for 38 CSF samples (paired t-test on log-transformed values, p = 0.145).

The stability of the total kit after a storage period of more than 2 years for different production runs was fully documented. Regression analysis on log-transformed values of concentrations for control samples did not show a statistically significant difference in P-tau$_{181P}$ concentrations as a function of storage time (kit stability). In addition, the P-tau$_{181P}$ concentration in CSF (n = 2) stored at −80°C did not change over a

period of more than 2 years, suggesting the possibility of storing CSF samples for extended periods. The CV for these samples amounted to 9.6% and 8.1% (n = 19 test runs) (sample stability).

Concentrated stock solutions of calibrator peptides were stable for more than 2 years when stored at 2–8°C (as compared to storage at −20°C). A Wilcoxon matched-pair test revealed no statistical differences between concentrations determined with the frozen or unfrozen calibrator peptides for three control samples. Working solutions of calibrators were repeatedly frozen and thawed (up to 10 times). Regression analysis on log-converted concentrations revealed a non-significant decrease for concentrations of 15.6 pg/mL ($p = 0.129$), 62.5 pg/mL ($p = 0.972$) and 125 pg/mL ($p = 0.069$), and a significant decrease for 31.25 pg/mL ($p = 0.003$). A Kruskal-Wallis test revealed that the significant difference for the 31.25 pg/mL was solely due to the values obtained after ten freeze/thaw cycles. It is essential to limit the number of freezing steps for the calibrators.

Analytical performance

Curve fitting Selection of the model for curve fitting was based on experiments with more than ten calibrators over a wide concentration range. The sigmoidal (four-parameter logistic) curves selected provided excellent fits ($R^2 = 0.99$ or higher) and residuals. There was an acceptable linear relation between the concentrations predicted from the calibrator curve or from an independent set of samples and the known concentrations over the complete calibration range of 15.6–500 pg/mL.

Analytical selectivity and specificity The AT270 mAb recognized tau phosphorylated at threonine 181. No reactivity was obtained using non-phosphorylated peptides or tau phosphorylated at other sites (data not shown). No cross-reactivity with recombinant tau was obtained in quantified CSF in concentration ranges corresponding to healthy and diseased subjects.

CSF handling procedures The inclusion of several freeze/thaw cycles (n = 5), as well as evaluation of the need for a centrifugation step before storage of the CSF, did not result in a significant effect on P-tau$_{181P}$ concentrations (data not shown). The effect of storage of CSF for a limited period is shown in Figure 1. No significant difference in P-tau$_{181P}$ concentration was noted between CSF samples stored for 3 days at room temperature and samples frozen immediately after collection ($p = 0.913$).

Linearity Three different pairs of CSF samples, containing high and low concentrations of P-tau$_{181P}$, were combined in different ratios. When the calculated concentrations were plotted against predicted concentrations, no deviation from the bisector could be observed, indicating the absence of matrix effects under the experimental conditions (Figure 2).

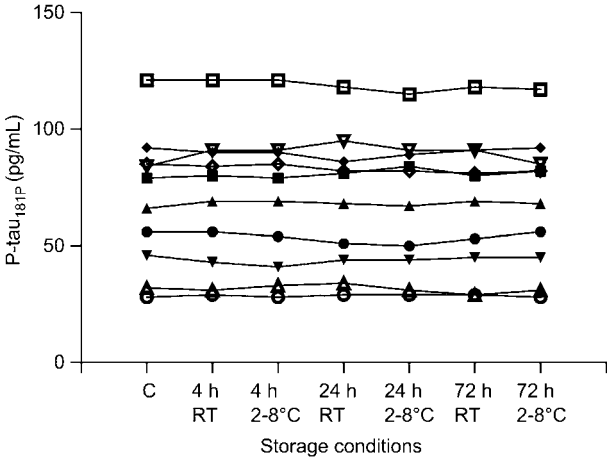

Figure 1 Storage and freezing. Effect of storage conditions of CSF samples on P-tau$_{181P}$ concentrations. C, stored immediately at −80°C; RT, room temperature, defined as 20–25°C.

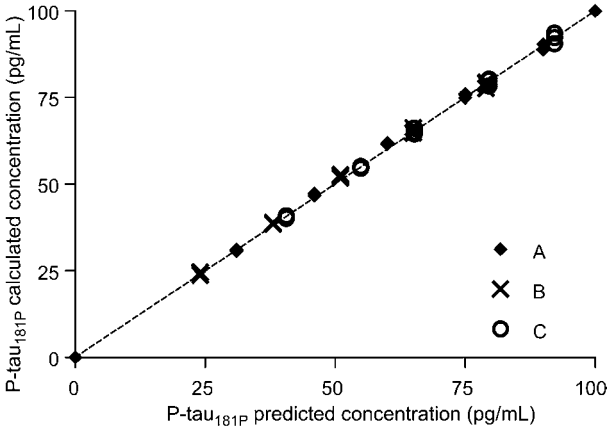

Figure 2 Assay linearity: combinations of CSF samples with high and low concentrations of P-tau$_{181P}$. A, B, C are different CSF combinations.

Repeatability and reproducibility The inter- and intra-lot variability was tested for low-, medium-, and high-reacting control samples. Results are shown in Table 2. All CV values (within or between production runs) were lower than 8.5%.

Assay sensitivity and precision Twelve samples were tested using eight replicates on three plates from five different production runs. The LOD determined was 10–15 pg/mL and LLOQ was 22–28 pg/mL. Figure 3A shows the precision profile of the pooled data for all the plates. Figure 3B shows the method performance chart (bias as a function of CV). The CV was less than 10% for levels between 15.6 and 300

Table 2 Repeatability (r) and reproducibility (R).

Sample	Intra-lot CV, %		Inter-lot CV, %	
	r	R	r	R
Low reactivity	3.1	7.1	2.7	8.0
Medium reactivity	2.8	6.5	1.4	5.1
High reactivity	2.5	8.4	2.5	3.2

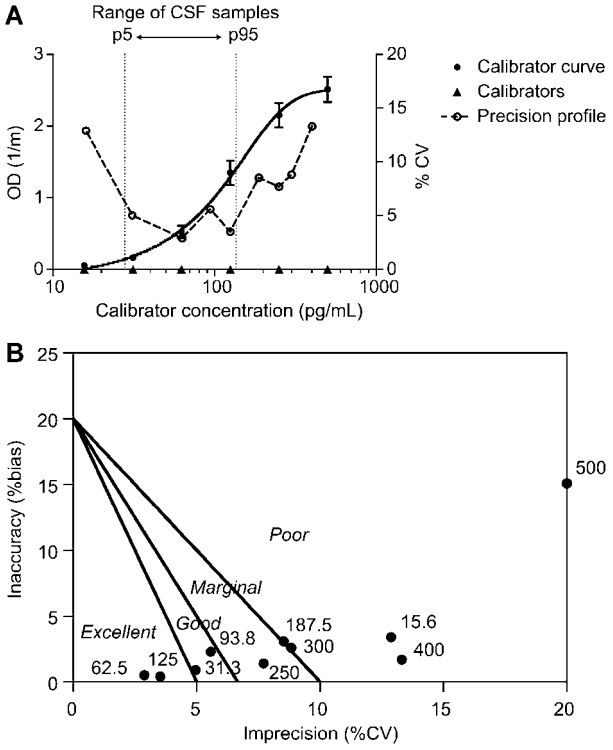

Figure 3 (A) Calibrator curve and precision profiles for calibrators of the INNOTEST® PHOSPHO-TAU$_{(181P)}$ assay. Precision profile expressed as percentage CV (o), with calibration curve plotted on the same graph (•) as well as the concentration range expressed as 5th (p5) and 95th percentiles (p95) for CSF samples included in the clinical study. (B) Method performance chart, established with calibrator peptides, for the INNOTEST® PHOSPHO-TAU$_{(181P)}$ assay. Values represent concentrations of calibrator peptides.

pg/mL. The higher calibrator concentrations showed lower precision. Bias was below 5% for concentrations between 15.6 and 300 pg/mL. Bias may exceed 10% for the highest concentrations. Good to excellent performance was obtained between LLOQ and ULOQ or higher. The validated range was estimated to be between 25 and 250 pg/mL.

Test robustness Different combinations of boundary assay parameters resulted in clear differences in OD output for control samples. However, at the level of P-tau$_{181P}$ concentrations, differences in boundary assay parameters (see also Table 1) remained within the range of random test variation. It was also noted that for the two samples in the linear test range, the CV was less than 10% in all assay parameter combinations tested (data not shown).

Clinical utility

Sample disposition The study was conducted between March and May 2002. Initially, 223 CSF samples were included in the study (Figure 4). Ten samples were excluded from statistical analysis since they did not meet the inclusion criteria. Statistical analysis revealed a center effect, with P-tau$_{181P}$ values generally lower for center 04. Criteria for pooled analysis were defined prospectively in the protocol. As a result, P-tau$_{181P}$ data from center 04 were omitted in the calculations for cutoff value, sensitivity, and specificity.

Diagnostic characteristics Demographic data and CSF concentrations of P-tau$_{181P}$, A$\beta_{1\text{-}42}$, and T-tau in the three diagnostic groups are described in Table 3. Independent of the center effect mentioned, significantly higher P-tau$_{181P}$ levels were observed for AD compared to DLB ($p<0.001$) and CS ($p<0.001$). No significant difference was observed between the DLB and CS groups ($p>0.05$). For A$\beta_{1\text{-}42}$ (data from five centers), significantly lower levels were observed for AD compared to CS ($p<0.001$) and for DLB compared to CS ($p<0.001$). No significant difference was observed between DLB and AD groups ($p=0.522$). For T-tau (data from five centers), significantly higher levels were observed for AD compared to CS ($p<0.001$) and DLB ($p<0.001$), while no differences were seen between CS and DLB ($p=0.077$).

In the present study, P-tau$_{181P}$ concentrations did not differ as a function of gender. No significant cor-

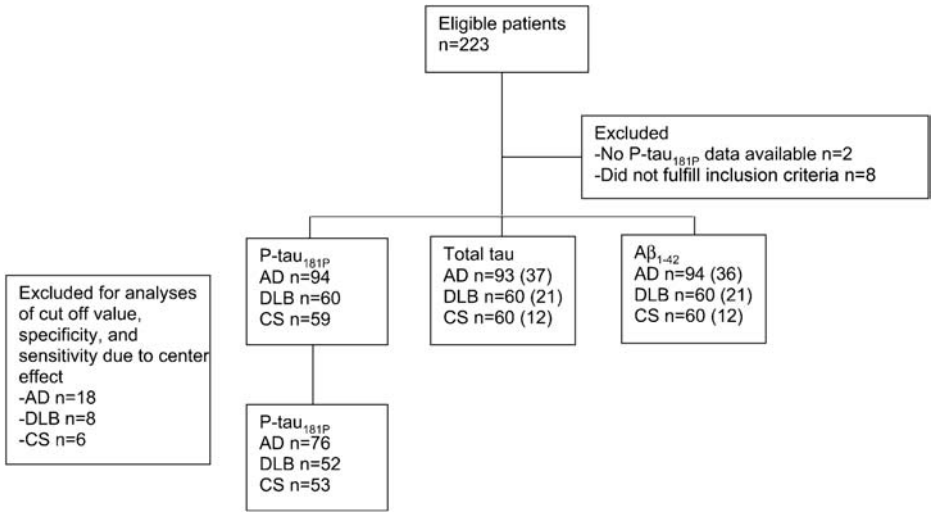

Figure 4 Summary of patient samples. Numbers in parentheses are historical results. AD, Alzheimer's disease; DLB, dementia with Lewy bodies; CS, control subjects.

Table 3 Demographic and biomarker data for patients enrolled in the study.

Patient data	AD	DLB	CS
Number of patients (M/F)	94 (55/39)	60 (21/39)	59 (35/24)
Age, years	67.5 (60.3, 77.0)	74.0 (71.0, 78.0) (*1*)	64.0 (58.8, 73.0)
MMSE	21.0 (15.0, 24.0) (*2*)	20.5 (17.0, 25.0) (*2*)	29.0 (28.3, 29.8) (*49*)
ApoE			
Missing data	27	22	32
No E4	26	20	18
One E4	32	18	9
Two E4	9	0	0
Biomarker levels, pg/mL			
P-Tau$_{181P}$	83.5 (54.3, 106.4)	47.6 (34.0, 58.0)	42.4 (34.3, 53.5) (*1*)
T-tau	610.0 (416.1, 868.5) (*1*)	271.5 (200.6, 356.7)	175.5 (127.7, 289.6)
A$\beta_{1\text{-}42}$	377.5 (285.0, 494.8)	433.9 (337.7, 532.3)	702.9 (550.9, 830.0)

AD, Alzheimer's disease; DLB, dementia with Lewy bodies; CS, control subjects. Results are expressed as median (25th, 75th percentiles). The number of missing values is shown in italics.

Table 4 Comparison of the AUC for P-tau$_{181P}$, A$\beta_{1\text{-}42}$, and T-tau for discrimination of AD from CS or DLB.

	AD vs. CS	AD vs. DLB
P-tau$_{181P}$	0.867 ± 0.031 (76/53; $p < 0.001$)	0.821 ± 0.036 (76/51; $p = 0.780$)
A$\beta_{1—42}$	0.881 ± 0.033 (77/54; $p = 0.010$)	0.588 ± 0.052 (77/52; $p < 0.001$)
T-tau	0.908 ± 0.026 (77/54; $p = 0.021$)	0.798 ± 0.038 (77/52; $p = 0.512$)
Model: A$\beta_{42} = 240 + 1.18 \times$tau (14)	0.960 ± 0.017	0.815 ± 0.037

Results are expressed as AUC ± SE (number of patients in the analysis, AD/CS or AD/DLB), with significance as compared to the model described in reference (14). Data were obtained from four centers; only results for which all parameters were available were included in the study. Significance values between parameters are further explained in the Results section.

relation between P-tau$_{181P}$ levels and age was found in CS ($p = 0.091$), AD ($p = 0.852$) or DLB ($p = 0.125$). In the AD and DLB groups, there was no correlation between apolipoprotein E (apoE) genotype and P-tau$_{181P}$ levels ($p > 0.05$). Furthermore, no significant correlation of P-tau$_{181P}$ was found with MMSE in AD ($r = 0.165$; $p = 0.160$) and DLB patients ($r = 0.085$; $p = 0.551$). In addition, no effect of blood-brain barrier deficits or gradient on P-tau$_{181P}$ concentrations was found (unpublished observations obtained from one center).

Classification A ROC curve describes true-positive vs. false-positive rates of a diagnostic test at different cutoff values. The area under the curve (AUC) is considered as a measure of the discriminative power of the test. Table 4 summarizes ROC curves for the differential diagnosis of AD-DLB and AD-CS. For AD vs. CS or DLB, the AUC obtained for T-tau or A$\beta_{1\text{-}42}$ was not significantly different between analyses with and without center 04 (data not shown). The AUC for T-tau was not statistically significantly different from the AUC for P-tau$_{181P}$ to discriminate AD vs. DLB ($p = 0.175$), but significantly higher than the AUC for P-tau$_{181P}$ to discriminate AD vs. CS ($p = 0.039$). The AUC for A$\beta_{1\text{-}42}$ was not significantly different from the AUC for P-tau$_{181P}$ to discriminate AD vs. CS ($p = 0.805$), but significantly lower than for P-tau$_{181P}$ to discriminate AD vs. DLB ($p < 0.001$).

To discriminate between AD and CS, P-tau$_{181P}$, T-tau, and A$\beta_{1\text{-}42}$ were entered into different algorithms, including a previously published model combining T-tau and A$\beta_{1\text{-}42}$ (15). Using discriminant analysis, a classification tree, or logistic regression, P-tau$_{181P}$ was removed from this model and diagnostic question, as it did not significantly contribute to the discrimination in this model. As a consequence, the center effect for P-tau$_{181P}$ was not relevant in this context. The resulting sensitivity, specificity, and overall accuracy of the models are shown in Table 5. To discriminate AD from CS, the AUC is significantly augmented using the previously described discrimination

Table 5 Sensitivity, specificity, and overall accuracy for discrimination of AD from CS.

	Sensitivity, %	Specificity, %	Overall accuracy, %
T-tau and A$\beta_{1\text{-}42}$			
Discrimination line (14)	92.3	83.1	88.7
Discriminant analysis	88	93	90
Discriminant analysis (log transformed values)	89	93	91
Classification tree	94.6	86.4	91.4
Logistic regression	91.3	89.8	90.7
P-tau$_{181P}$ and A$\beta_{1\text{-}42}$			
Discriminant analysis	84	85	84
Discriminant analysis (log transformed values)	85	86	85

Table 6 Sensitivity, specificity, and overall accuracy for discrimination of AD from DLB.

	Sensitivity, %	Specificity, %	Overall accuracy, %
Discriminant analysis (P-tau$_{181P}$)	74	85	78
Discriminant analysis (log transformed values)	82	81	81
Classification tree (P-tau$_{181P}$)	80	79	80

line (14) compared to T-tau ($p=0.021$), P-tau$_{181P}$ ($p<0.001$) or Aβ_{1-42} ($p=0.010$) analyzed individually.

By analogy, P-tau$_{181P}$, T-tau, and Aβ_{1-42} were entered into algorithms to discriminate AD from DLB. The point on the P-tau$_{181P}$ ROC curve at which sensitivity and specificity are maximal for differentiation of AD from DLB was determined as 61 pg/mL. Using this cutoff, a sensitivity of 80% and specificity of 79% were obtained (Table 6).

However, in the discriminant analysis model for AD vs. DLB, it is important to note that P-tau$_{181P}$ alone is the most important variable in the discrimination between AD from DLB. P-tau$_{181P}$ and Aβ_{1-42} remained in the discriminant analysis log-transform model, whereas T-tau was rejected. The most important variable in the classification tree was P-tau$_{181P}$, resulting in a cutoff value of 61 pg/mL. P-tau$_{181P}$ was also the most important variable in the logistic regression model. The cutoff value from the logistic regression model was calculated as 59.4 pg/mL. The resulting sensitivity, specificity, and overall accuracy of the different models are shown in Table 6.

Discussion

We describe here the analytical and clinical performance of the INNOTEST® PHOSPHO-TAU$_{(181P)}$ ELISA for quantification of P-tau$_{181P}$ in CSF. Both phosphorylated and non-phosphorylated tau isoforms have been identified in CSF, fetal human brain, and adult human brain tissue (16, 17). We found that the quantification of P-tau$_{181P}$ levels in CSF, using a combination of mAbs AT270 and HT7, was specific; reactivity towards other phosphorylated epitopes and non-phosphorylated tau protein was absent, confirming previously published characterization data (10, 16, 18). Linearity experiments pointed to a high accuracy level when using undiluted CSF samples. When reproducibility and repeatability experiments were performed (with operators, lot numbers, or test runs as variables), the CV for samples in the assay range were below 15%. Notwithstanding the validated robust assay format, the stability of the kit for at least 13 months (as stated in the kit insert), the relative stability of P-tau$_{181P}$ in CSF, and the possibility to freeze/thaw calibrator peptides, it is still important to carry out on-site training to obtain the best assay performance. This could avoid center effect(s) as observed in the present study. The production process resulted in a real-time stability of at least 13 months. Extension of this shelf life in the future will be an advantage in clinical trials, since clinical trials sometimes run for more than 2 years before finalization; kits from a single lot could be used to analyze all samples in a specific study. In addition, different CSF handling protocols were evaluated. No effect on P-tau$_{181P}$ concentrations was noted after centrifugation or storage for up to 3 days at 25°C before freezing at −80°C. However, repeated freezing and thawing of CSF is not recommended, especially when P-tau$_{181P}$ is to be analyzed in combination with other biomarkers, such as Aβ_{1-42} (9).

The results of the clinical study extend a previous report (5) on the value of P-tau$_{181P}$ testing in CSF for discriminating AD from DLB. An overall accuracy level of 80% was observed, which may be difficult to improve on by single-parameter analysis, considering the level of accuracy of clinical diagnosis compared with postmortem findings. The present study also confirmed the reported discrimination line (15) based on the combination of CSF T-tau and Aβ_{1-42} for discriminating AD from CS. The use of P-tau$_{181P}$ could not improve this result. However, P-tau$_{181P}$ was shown to be the most important variable among the different biomarkers in discriminating between AD and DLB, although AUC differences for P-tau$_{181P}$ compared to T-tau did not reach statistical significance. Notably, when P-tau$_{181P}$ was used to classify patients instead of T-tau, six additional patients were correctly classified as DLB, whereas Aβ_{1-42} was shown to have no discriminatory power in distinguishing AD from DLB. Moreover, using different statistical models based on the three biomarkers, P-tau$_{181P}$ was consistently identified as the most important marker in the discrimination of AD from DLB.

In the study of Mollenhauer et al. (19), significantly different values were obtained for the comparison between AD and DLB for T-tau and the T-tau/P-tau$_{181P}$ ratio, but not for Aβ_{1-42}, P-tau$_{181P}$, or the mixed Aβ_{1-42}/Aβ_{40} ratio. Notwithstanding the fact that in this study values for P-tau$_{181P}$ were lower, albeit not significantly, in the DLB group compared to AD, differences could be explained by the sample size, patients enrolled in the study and/or the homogeneity of the study group. These results will have to be confirmed in larger studies.

A number of sandwich immunoassays that target different phosphorylated epitopes of tau, including Thr181 and Thr231, Thr231 and Ser235, Ser199, Thr231, Thr181, and Ser396, and Ser404 have been developed (20). Consistently increased levels of P-tau in CSF from patients with AD have been found using these assays (3). In a study directly comparing the diagnostic performance of P-tau$_{181P}$, P-tau$_{199}$, and P-tau$_{231}$ in the same sets of patients, all three assays performed equally well in the discrimination of AD patients from non-demented subjects. Only minor differences were found when group separation was maximized between AD and FTD using P-tau$_{231}$ and between AD and DLB using P-tau$_{181P}$ controls (21).

Furthermore, there was no difference in the prediction of cognitive decline using each of the P-tau epitopes in patients with mild cognitive impairment (22). The present study confirms findings published in previous reports that quantification of P-tau could be a useful tool for differential diagnosis of neurodegenerative disorders (5, 21, 23).

Numerous published reports have already mentioned the need for a multiparametric approach in the domain of AD diagnosis. As in other areas of medicine, CSF markers for AD may not be used in isolation, but can provide added value to the information obtained from clinical examination, brain-imaging techniques (e.g., single photon emission CT and MRI), and routine CSF biochemical analysis. It has already been concluded that T-tau, Aβ_{42}, and P-tau$_{181P}$, when used as adjuncts to clinical diagnosis, have the potential to help differentiate AD from some difficult differential diagnoses (e.g., normal aging, depressive pseudo-dementia, Parkinson's disease, progressive supranuclear palsy, and alcoholic dementia) (24) and amyotrophic lateral sclerosis (25). The present study confirms that better diagnostic performance compared to single-parameter analysis can be obtained for the discrimination of CS and AD using T-tau and Aβ_{1-42} (15). The P-tau$_{181P}$/T-tau ratio can discriminate patients with Creutzfeldt-Jakob disease from those with other neurological disorders, including AD and FTD (26, 27). Similar cutoff values were published for discrimination of AD patients from controls by Lewczuk et al. (28). Higher diagnostic accuracy was reported for the combination of low CSF Aβ_{1-42} and high P-tau$_{181P}$ in differentiating early-onset AD from FTD (6) and for differentiation of AD and healthy controls (29). Identical results with respect to Aβ_{1-42} and T-tau for comparison between CS, AD, and DLB have been published recently (30).

Individual markers have been demonstrated to be of potential clinical use in the identification of early AD. Further improvement will likely be achieved by the creation of a complex model that includes several biomarkers to adequately describe this phenomenon. Published results on the diagnostic capacity of CSF markers come from studies of clinically diagnosed patients. This means that the diagnostic performance of CSF markers cannot be higher than the accuracy of the clinical diagnostic criteria used. Owing to this diagnostic uncertainty and the heterogeneity of the disease process, it may well be beneficial to select, validate, and implement an algorithm that is based on a combination of biomarkers to reduce the overlap between different diagnostic groups.

The data available at present must be used with caution when building ''optimistic'' models or algorithms to estimate the performance in terms of clinical accuracy. The clinical accuracy levels in the present paper are mentioned as tentative estimates and not as a final conclusion for a particular application. The model should be validated and confirmed using a larger number of samples.

A limitation of the present study is, of course, the small cohort with the risk of selection bias and insufficient power to draw firm conclusions. The concentration and cutoff values provided can be used for guidance. However, given the center effect observed, laboratories using this marker in a clinical setting should first establish local reference ranges.

In conclusion, these data show that the INNOTEST® PHOSPHO-TAU$_{(181P)}$ is a robust and reliable assay that can potentially be used in a diagnostic context. It will be important to accumulate data from longitudinal studies to interpret results of the biomarkers together with imaging and neuropsychological markers.

Acknowledgements

The authors thank Pierre Grognet, Els Van Kerschaever, Leentje Demeyer, and Cristiano Spaccatini for CSF analyses; Geert De Meyer for statistical analysis; and Jan Brissink, Fred Shapiro, and Anne Farmer for manuscript editing. The study was supported in part by IWT (Instituut voor aanmoediging van Innovatie door Wetenschap en Technologie) Grant 990220.

References

1. Wiltfang J, Esselman H, Maler JM, Bleich S, Hüther G, Kornhuber J. Molecular biology of Alzheimer's dementia and its clinical relevance to early diagnosis and new therapeutic strategies. Gerontology 2001;47:65–71.
2. Buée L, Bussière T, Buée-Scherrer V, Delacourte A, Hof PR. Tau protein isoforms, phosphorylation and role in neurodegenerative disorders. Brain Res Rev 2000;33: 95–130.
3. Blennow K, Hampel H. CSF markers for incipient Alzheimer's disease. Lancet Neurol 2003;2:605–13.
4. Consensus report of the working group on molecular and biochemical markers of Alzheimer's disease. The Ronald and Nancy Reagan Research Institute of the Alzheimer's Association and the National Institute on Aging Working Group. Neurobiol Aging 1998;19:109–16 (Erratum in Neurobiol Aging 1998;19:285).
5. Vanmechelen E, Van Kerschaver E, Blennow K, De Deyn PP, Galasko D, Parnetti L, et al. CSF-Phospho-tau (181P) as a promising marker for discriminating Alzheimer's disease from dementia with Lewy bodies. In: Iqbal K, Sisodia SS, Winblad B, editors. Alzheimer's disease: advances in etiology, pathogenesis and therapeutics. New York: Wiley, 2001;285–91.
6. Schoonenboom NS, Pijnenburg YA, Mulder C, Rosso SM, Van Elk EJ, van Kamp GJ, et al. Amyloid $\beta_{(1-42)}$ and phosphorylated tau in CSF as markers for early-onset Alzheimer disease. Neurology 2004;62:1580–4.
7. Goedert M, Spillantini MG, Jakes R, Rutherford D, Crowther RA. Multiple isoforms of human microtubule-associated protein tau: sequences and localization in neurofibrillary tangles of Alzheimer's disease. Neuron 1989;3:519–26.
8. Vandermeeren M, Mercken M, Vanmechelen E, Six J, Van de Voorde A, Martin J-J, et al. Detection of τ proteins in normal and Alzheimer's disease cerebrospinal fluid with a sensitive sandwich enzyme-linked immunosorbent assay. J Neurochem 1993;61:1828–34.
9. Vanderstichele H, Van Kerschaver E, Hesse C, Davidsson P, Buyse M-A, Andreasen N, et al. Standardization of measurement of β-amyloid$_{(1-42)}$ in cerebrospinal fluid and plasma. Amyloid 2000;7:245–58.
10. Vanmechelen E, Vanderstichele H, Davidsson P, Van Kerschaver E, Van Der Perre B, Sjögren M, et al. Quan-

tification of tau phosphorylated at threonine 181 in human cerebrospinal fluid: a sandwich ELISA with a synthetic phosphopeptide for standardization. Neurosci Lett 2000;285:49–52.
11. ISO 5725–2: Accuracy (trueness and precision) of measurement methods and results - Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method, Geneva: International Organization for Standardization, 1994.
12. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology 1984;34: 939–44.
13. McKeith IG, Galasko D, Kosaka K, Perry EK, Dickson DW, Hansen LA, et al. Consensus guidelines for the clinical and pathologic diagnosis of dementia with Lewy bodies (DLB): report of the consortium on DLB international workshop. Neurology 1996;47:1113–24.
14. Folstein MF, Folstein SE, McHugh PR. ''Mini-mental state''. A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12: 189–98.
15. Hulstaert F, Blennow K, Ivanoiu A, Schoonderwaldt HC, Riemenschneider M, De Deyn PP, et al. Improved discrimination of AD patients using β-amyloid$_{(1-42)}$ and tau levels in CSF. Neurology 1999;52:1555–62.
16. Goedert M, Jakes R, Crowther RA, Cohen P, Vanmechelen E, Vandermeeren M, et al. Epitope mapping of monoclonal antibodies to the paired helical filaments of Alzheimer's disease: identification of phosphorylation sites in tau protein. Biochem J 1994;301:871–7.
17. Sjögren M, Davidsson P, Tullberg M, Minthon L, Wallin A, Wikkelso C, et al. Both total and phosphorylated tau are increased in Alzheimer's disease. J Neurol Neurosurg Psychiatry 2001;70:624–30.
18. Olsson A, Vanderstichele H, Andreasen N, De Meyer G, Wallin A, Holmberg B, et al. Simultaneous measurement of β-amyloid$_{(1-42)}$, total tau, and phosphorylated tau (Thr^{181}) in cerebrospinal fluid by the xMAP Technology. Clin Chem 2005;51:336–45.
19. Mollenhauer B, Bibl M, Wiltfang J, Steinacker P, Ciesielczyk B, Neubert K, et al. Total tau protein, posphorylated tau (181p) protein, β-amyloid$_{1-42}$, and β-amyloid$_{1-40}$ in cerebrospinal fluid of patients with dementia with Lewy bodies. Clin Chem Lab Med 2006;44:192–5.
20. Blennow K. Cerebrospinal fluid protein biomarkers for Alzheimer's disease. NeuroRx 2004;1:213–25.
21. 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.
22. Buerger K, Ewers M, Andreasen N, Zinkowski R, Ishiguro K, Vanmechelen E, et al. Phosphorylated tau predicts rate of cognitive decline in MCI subjects – a comparative CSF study. Neurology 2005;65:1502–3.
23. Buerger K, Zinkowski R, Teipel SJ, Tapiola T, Arai H, Blennow K, et al. Differential diagnosis of Alzheimer disease with cerebrospinal fluid levels of tau protein phosphorylated at threonine 231. Arch Neurol 2002;59:1267–72.
24. Nägga K, Gottfries J, Blennow K, Marcusson J. Cerebrospinal fluid phospho-tau, total tau and beta-amyloid(1-42) in the differentiation between Alzheimer's disease and vascular dementia. Dement Geriatr Cogn Disord 2002;14:183–90.
25. Jimenez-Jimenez FJ, Hernanz A, Medina-Acebron S, de Bustos F, Zurdo JM, Alonso H, et al. Tau protein concentrations in cerebrospinal fluid of patients with amyotrophic lateral sclerosis. Acta Neurol Scand 2005;111: 114–7.
26. Riemenschneider M, Wagenpfeil S, Vanderstichele H, Otto M, Wiltfang J, Kretzschmar H, et al. Phospho-tau/total tau ratio in cerebrospinal fluid discriminates Creutzfeldt-Jakob disease from other dementias. Mol Psychiatry 2003;8:343–7.
27. Zetterberg H, Blennow K. Elevated total tau/phospho-tau ratio in autopsy-proven Creutzfeldt-Jakob disease with negative 14-3-3 test results. Neurol Sci 2004;25:301–2.
28. Lewczuk P, Esselmann H, Bibl M, Beck G, Maler JM, Otto M, et al. Tau protein phosphorylated at threonine 181 in CSF as a neurochemical biomarker in Alzheimer's disease: original data and review of the literature. J Mol Neurosci 2004;23:115–22.
29. Maddalena A, Papassotiropoulos A, Müller-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 β-amyloid peptide$_{42}$. Arch Neurol 2003;60:1202–6.
30. Mollenhauer B, Cepek L, Bibl M, Wiltfang J, Schulz-Schaeffer WJ, Ciesielczyk B, et al. Tau protein, Aβ42 and S-100B protein in cerebrospinal fluid of patients with dementia with Lewy bodies. Dement Geriatr Cogn Disord 2005;19:164–70.

Received May 13, 2006, accepted September 14, 2006