Published in final edited form as:

Am J Clin Pathol. 2008 April ; 129(4): 526–529. doi:10.1309/W01Y0B808EMEH12L.

CSF Multianalyte Profile Distinguishes Alzheimer and Parkinson Diseases

Jing Zhang, MD, PhD¹, Izabela Sokal, PhD¹, Elaine R. Peskind, MD^{2,3}, Joseph F. Quinn, MD⁴, Joseph Jankovic, MD⁵, Christopher Kenney, MD⁵, Kathryn A. Chung, MD, PhD⁴, Steven P. Millard, PhD³, John G. Nutt, MD⁴, and Thomas J. Montine, MD, PhD¹

¹Department of Pathology, University of Washington School of Medicine, Seattle ²Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine, Seattle ³Mental Illness Research, Education, and Clinical Center, Veterans Affairs Puget Sound Health Care System, Seattle ⁴Department of Neurology, Oregon Health and Science University, Portland ⁵Parkinson's Disease Center and Movement Disorders Clinic, Department of Neurology, Baylor College of Medicine, Houston, TX.

Abstract

The therapeutic imperative for Alzheimer disease (AD) and Parkinson disease (PD) calls for discovery and validation of biomarkers. Increased cerebrospinal fluid (CSF) τ and decreased amyloid (A) β_{42} have been validated as biomarkers of AD. In contrast, there is no validated CSF biomarker for PD. We validated our proteomics-discovered multianalyte profile (MAP) in CSF from 95 control subjects, 48 patients with probable AD, and 40 patients with probable PD. An optimal 8-member MAP agreed with expert diagnosis for 90 control subjects (95%), 36 patients with probable AD (75%), and 38 patients with probable PD (95%). This MAP consisted of the following (in decreasing order of contribution): τ , brain-derived neurotrophic factor, interleukin 8, A β_{42} , β_2 -microglobulin, vitamin D binding protein, apolipoprotein (apo) AII, and apoE. This first large-scale validation of a proteomic-discovered MAP suggests a panel of 8 CSF proteins that are highly effective at identifying PD and moderately effective at identifying AD.

Keywords

Cerebrospinal fluid; Alzheimer disease; Parkinson disease; Biomarkers; Analyte profile; Random forest algorithm

Alzheimer disease (AD) and Parkinson disease (PD) are major public health problems. Key to the effort to obtain new therapeutics will be novel biomarkers to aid in diagnosis, identify subsets of patients, and objectively monitor progression and response to treatment. Several discovery proteomic studies of human cerebrospinal fluid (CSF) have been reported using relatively small numbers of patients with AD compared with control subjects without dementia.^{1–6} Few proteomic studies have attempted to discriminate AD from PD or other neurodegenerative diseases, and fewer still have used CSF samples from patients who subsequently had autopsy verification of clinical diagnosis.^{1,2} The extent of confirmation of these potential biomarkers discovered by proteomic techniques has varied among studies but, again, has been limited to relatively few samples (10 to 20 per group) of one disease group vs

control subjects.^{1,4,5} There has been no report of large-scale validation of these potential CSF biomarkers for AD and PD.

We report results of a large-scale validation of the best 7 potential candidates from our previous CSF proteomics studies for discriminating AD and PD^{1,4}: brain-derived neurotrophic factor (BDNF), interleukin (IL)-8, vitamin D binding protein (VDBP), β_2 -microglobulin, haptoglobin, apolipoprotein (apo) AII, and apoE, plus apoAI that was part of a multiplexed kit. In addition to these, we measured CSF levels of 2 validated biomarkers for AD (amyloid [A] β_{42} and τ) in 202 patients using X-multianalyte profile (MAP) technology to permit facile translation to clinical laboratories to improve the diagnosis of common neurodegenerative diseases, especially in settings without ready access to subspecialty-trained clinicians.

Materials and Methods

Criteria for Selection of Patients and Control Subjects

Human subjects divisions at each institution approved this study. Following informed consent, all subjects underwent extensive evaluation that consisted of medical history, family history, physical and neurologic examinations by clinicians who specialize in movement disorders or dementia, laboratory tests, and neuropsychological assessment; information was obtained from control subjects or from informants for patients. Control subjects were compensated community volunteers in good health with no signs or symptoms suggesting cognitive decline or neurologic disease. All patients were participants in research clinical cores at our respective institutions. Clinical diagnoses were made by board-certified neurologists or geriatric psychiatrists who are faculty in National Institutes of Health–funded Alzheimer disease centers or board-certified neurologists who are movement disorder specialists at academic medical centers. Clinical diagnoses of mild cognitive impairment (MCI), AD, PD, and frontotemporal dementia (FTD) were made exactly following well-established consensus criteria.^{7–10}

All CSF samples obtained by intra vitam lumbar puncture between December 2001 and January 2006 were eligible except for samples used in previous discovery experiments. Inclusion criteria were CBC count and serum electrolyte, serum urea nitrogen, creatinine, glucose, vitamin B₁₂, and thyroid stimulating hormone results within normal limits. Exclusion criteria included heavy cigarette smoking (>10 pack-years) and alcohol use other than social. Any psychotherapeutic medication use was an exclusion criterion for control subjects, and any psychotherapeutic use other than for treatment of neurodegenerative disease was an exclusion criterion for patient groups. Because the youngest patient with neurodegenerative disease was 37 years old, we limited control subjects to people 35 years or older. No person whose CSF sample was investigated in this study has died and undergone autopsy at this time.

Quality Control for CSF

CSF was obtained by lumbar puncture in the morning. CSF was collected at 3 study sites using an identical method,¹¹ stored at -80°C immediately after collection, and never thawed before analysis. All CSF samples archived from patients or control subjects that met these criteria were used unless visibly contaminated by blood. This is a departure from our discovery experiments in which we strictly controlled for blood contamination^{1,4}; however, in the present study, we were trying to determine performance characteristics of a MAP profile in a more clinically relevant setting.

Analysis of CSF

Multiplex analysis of 10 CSF proteins was performed using immunobead-based multiplex assays purchased from 2 sources: Biosource Division of Invitrogen, Camarillo, CA (τ , A β_{42} , BDNF, VDBP, haptoglobin, and β_2 -microglobulin), and Linco Research, St Charles, MO

(apoE, apoAI, apoAII, and IL-8). The samples and standards were run according to the manufacturer's instructions using a LiquiChip Workstation (Qiagen, Valencia, CA); values were calculated by interpolation from standard curves.

Statistical Analysis

Data were analyzed by using Breiman's random forest (RF) algorithm for classification (available at http://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm),¹² available as part of the R software suite (<http://www.r-project.org/>).¹³ The RF algorithm is a means of unbiased clustering based on the construction of many decision trees that estimate the importance of variables to the classification of subjects. We used the 10 proteins plus age and sex as input variables, the default value of *mtry* (14; the number of variables randomly chosen at each split in a tree), and the default value of *ntree* (500; the number of trees grown). The optimal number of variables was determined with the *varSelRF* algorithm.¹⁴

Results

Demographic characteristics of the subjects whose CSF samples were used in this study are given in Table 1. A summary of results from CSF measurements for 183 subjects in the control, AD, and PD groups is given in Table 2. Patients with PD had mild to moderate impairment with a median Hoehn and Yahr Staging Score of 2 (range, 1–3).¹⁵ Patients with AD had mild to moderate cognitive impairment with a median clinical dementia rating of 1 (range, 0.5–2).¹⁶ When considering each measurement separately, as expected from numerous previous studies, τ was increased and $A\beta_{42}$ decreased when comparing control with AD but not control with PD groups (reviewed by Galasko¹⁷). BDNF, IL-8, VDBP, apoAII, and apoE were significantly different between the control and both neurodegenerative groups but not between neurodegenerative groups; of these, IL-8 and VDBP were increased in neurodegenerative diseases, whereas the remainder were decreased compared with control samples, again consistent with our previous quantitative proteomic studies.^{1,4} β_2 -microglobulin was significantly increased only between the control and PD groups, whereas apoAI and haptoglobin were not different among the 3 groups. Of the proteins identified in previous proteomic work as potential biomarkers for AD or PD, all but haptoglobin were confirmed in the present study. The AD group was older on average, and the control group had more women.

These 12 variables (10 CSF proteins, age, and sex) were entered into an RF analysis.^{12–14} Table 2 shows the 12 variables in decreasing order of relative contribution to discriminating among the 3 groups, as measured by mean decrease in accuracy. For each tree, the prediction accuracy for the cases not used to construct the tree was recorded, and this process repeated after permuting the values of each predictor variable. The mean decrease in accuracy was computed by averaging the differences between these 2 accuracy estimates for all trees and normalizing by the SE.

Table 3 lists the RF classification rates for each group using all 12 variables or subsets of the top-ranked variables. The top 8 variables offered optimal classification.¹⁴ It is important to note that these top 8 excluded age, sex, apoAI (measured simply because it was contained in a prefabricated kit), and haptoglobin (the one unconfirmed protein). By using the top 8 variables for classification, all misclassified control subjects were incorrectly grouped as having AD, the 12 misclassified patients with AD were incorrectly grouped as 10 control subjects and 2 patients with PD, and the 2 misclassified patients with PD were incorrectly grouped as 1 control subject and 1 patient with AD.

In addition to the 183 control, AD, and PD CSF samples, we also analyzed available CSF samples that met our inclusion and exclusion criteria from 12 patients with MCI and 7 with FTD. By using all 12 variables or only variables 1 through 8, 9 of 12 patients with MCI were

classified as control subjects and the other 3 were classified as having AD; none were classified as having PD. Again, by using all 12 or only the top 8 variables, 5 of 7 patients with FTD were classified as control subjects and the other 2 were classified as having AD; none were classified as having PD.

Discussion

We report on the first large-scale confirmation of CSF proteomic experiments that sought novel biomarkers for AD and PD. We assembled our MAP from candidate biomarkers in previous proteomic studies and the standard pair of τ and $A\beta_{42}$ and used Luminex technology (Austin, TX) because this is the likely platform for translation to clinical laboratories.^{1,4} Our results confirmed the usefulness of 6 of 7 proteins suggested as new CSF biomarkers by proteomic screens; haptoglobin failed to be validated. CSF τ and $A\beta_{42}$ showed the expected differences among groups, and the superfluous apoAI was nonsignificant. Of the validated proteins, only τ and $A\beta_{42}$ were significantly different between AD and PD, and the remainder were significantly different between the control group and the AD or PD group but not between AD and PD groups.

RF analysis was used to determine the optimal combination among these 10 CSF protein measurements, age, and sex for correctly classifying 183 CSF samples into one of the groups defined by extensive clinical testing by expert clinicians and neuropsychologists: control, AD, or PD. Optimal grouping was achieved with 8 CSF protein measurements but not age or sex. The rank order of contribution to correct classification (from most to least) for these 8 was τ , BDNF, IL-8, $A\beta_{42}$, β_2 -microglobulin, VDBP, apoAII, and apoE. It is perhaps not unexpected that each of these proteins already has been linked somehow to the pathogenesis of AD, PD, or both.

In addition to analysis of CSF samples from relatively large groups of patients with probable AD or PD, we also applied our MAP to smaller groups. Our MAP recognized the majority of FTD cases as different from AD and, therefore, when combined with the clinical diagnosis of dementia, may be useful in enriching patient populations for clinical trials or choosing appropriate treatments as they become available. Our MAP classified 3 of 12 patients with MCI as having AD. Because MCI is an enrichment for people with prodromal AD but also includes people with cognitive impairment due to other causes, it is not possible at this point to draw conclusions about the performance of our MAP in this group; time and much more analysis will tell if our MAP is useful in identifying people with prodromal AD or perhaps PD. Finally, emerging therapeutics for these diseases will require objective means to assess disease progression; ongoing studies will determine if some components of our MAP or other as yet only partially identified proteins¹⁸ may be helpful in this regard.

Acknowledgments

Supported by National Institutes of Health grant AG05136, the Michael J. Fox Foundation for Parkinson's Research, and the Buster and Nancy Alvord Endowment at the University of Washington, Seattle.

References

1. Abdi F, Quinn JF, Jankovic J, et al. Detection of biomarkers with a multiplex quantitative proteomic platform in cerebrospinal fluid of patients with neurodegenerative disorders. *J Alzheimers Dis* 2006;9:293–348. [PubMed: 16914840]
2. Castano EM, Roher AE, Esh CL, et al. Comparative proteomics of cerebrospinal fluid in neuropathologically-confirmed Alzheimer's disease and non-demented elderly subjects. *Neurol Res* 2006;28:155–163. [PubMed: 16551433]

3. Puchades M, Hansson SF, Nilsson CL, et al. Proteomic studies of potential cerebrospinal fluid protein markers for Alzheimer's disease. *Brain Res Mol Brain Res* 2003;118:140–146. [PubMed: 14559363]
4. Zhang J, Goodlett DR, Quinn JF, et al. Quantitative proteomics of cerebrospinal fluid from patients with Alzheimer disease. *J Alzheimers Dis* 2005;7:125–133. [PubMed: 15851850]
5. Finehout EJ, Franck Z, Choe LH, et al. Cerebrospinal fluid proteomic biomarkers for Alzheimer's disease. *Ann Neurol* 2007;61:120–129. [PubMed: 17167789]
6. Davidsson P, Westman-Brinkmalm A, Nilsson C, et al. Proteome analysis of cerebrospinal fluid proteins in Alzheimer patients. *Neuroreport* 2002;13:611–615. [PubMed: 11973456]
7. Petersen RC, Smith GE, Waring SC, et al. Mild cognitive impairment: clinical characterization and outcome. *Arch Neurol* 1999;56:303–308. [PubMed: 10190820]
8. McKhann G, Drachman D, Folstein M, et al. Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of the Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 1984;34:939–944. [PubMed: 6610841]
9. Gelb DJ, Oliver E, Gilman S. Diagnostic criteria for Parkinson disease. *Arch Neurol* 1999;56:33–39. [PubMed: 9923759]
10. Neary D, Snowden JS, Gustafson L, et al. Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. *Neurology* 1998;51:1546–1554. [PubMed: 9855500]
11. Peskind ER, Riekse R, Quinn JF, et al. Safety and acceptability of the research lumbar puncture. *Alzheimer Dis Assoc Disord* 2005;19:220–225. [PubMed: 16327349]
12. Breiman L. Random forests. *Machine Learning* 2001;45:5–32.
13. Liaw A, Wiener M. Classification and regression by randomForest. *R News* 2002;2:18–22.
14. Diaz-Uriarte R, Alvarez de Andrés S. Gene selection and classification of microarray data using random forest. *BMC Bioinformatics* 2006;7:3. [PubMed: 16398926]doi: 10.1186/1471-2105-7-3.
15. Goetz CG, Poewe W, Rascol O, et al. Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: status and recommendations. *Mov Disord* 2004;19:1020–1028. [PubMed: 15372591]
16. Morris JC. Clinical dementia rating: a reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. *Int Psychogeriatr* 1997;9:173–178. [PubMed: 9447441]
17. Galasko D. Biomarkers for Alzheimer's disease: clinical needs and application. *J Alzheimers Dis* 2005;8:339–346. [PubMed: 16556965]
18. Simonsen AH, McGuire J, Hansson O, et al. Novel panel of cerebrospinal fluid biomarkers for the prediction of progression to Alzheimer dementia in patients with mild cognitive impairment. *Arch Neurol* 2007;64:366–370. [PubMed: 17353378]

Table 1

Demographic Characteristics of Groups

	Control (n = 95)	Alzheimer Disease (n = 48)	Parkinson Disease (n = 40)	Mild Cognitive Impairment (n = 12)	Frontotemporal Dementia (n = 7)
Mean ± SD age (y)	63 ± 12	70 ± 9	59 ± 10	71 ± 12	65 ± 6
No. (%) female	51 (54)	19 (40)	9 (23)	4 (33)	3 (43)

Table 2

Cerebrospinal Measurements and RF Analysis*

	Control Group (n = 95)	AD Group (n = 48)	PD Group (n = 40)	RF Analysis	
				Variable	MDA
τ (pg/mL) ^{†,‡,§}	690.2 ± 39.1	1,425.0 ± 104.3	387.8 ± 46.9	1	2.27
BDNF (pg/mL) ^{†,‡}	241.0 ± 3.7	202.3 ± 4.5	184.0 ± 5.5	2	1.94
IL-8 (pg/mL) ^{†,‡}	16.4 ± 0.9	37.4 ± 3.5	36.3 ± 4.7	3	1.92
Amyloid β_{42} (pg/mL) ^{†,‡,§}	371.8 ± 20.5	268.5 ± 27.1	510.6 ± 27.8	4	1.82
β_2 -microglobulin (μg/mL) [‡]	1.1 ± 0.0	1.4 ± 0.1	1.6 ± 0.1	5	1.75
VDBP (μg/mL) ^{†,‡}	0.6 ± 0.0	1.1 ± 0.1	1.2 ± 0.1	6	1.75
ApoAII (μg/mL) ^{†,‡}	1.5 ± 0.1	0.9 ± 0.1	0.8 ± 0.2	7	1.62
ApoE (μg/mL) ^{†,‡}	3.9 ± 0.3	2.5 ± 0.2	2.3 ± 0.1	8	1.38
Mean ± SD age (y) ^{†,§}	63 ± 12	70 ± 9	59 ± 10	9	1.08
ApoAI (μg/mL)	2.6 ± 0.1	2.3 ± 0.2	2.4 ± 0.3	10	0.81
Haptoglobin (μg/mL)	1.5 ± 0.2	2.3 ± 0.6	3.8 ± 1.1	11	0.39
No. (%) female	51 (54)	19 (40)	9 (23)	12	0.13

apo, apolipoprotein; BDNF, brain-derived neurotrophic factor; IL, interleukin; MDA, mean decreased accuracy; PD, Parkinson disease; RF, random forest; VDBP, vitamin D binding protein.

* Data are given as mean ± SEM unless otherwise indicated. If nonparametric analysis of variance (Kruskal-Wallis test) resulted in $P < .01$, it was followed by Dunn corrected multiple comparisons with

[†] $P < .01$ for control vs AD groups

[‡] $P < .01$ for control vs PD groups

[§] $P < .01$ for AD vs PD groups. Results from RF analysis are the ranking of variables from most to least contribution to separating groups, MDA in grouping when one variable was excluded from analysis.

^{||} $P < .01$ by χ^2 test.

Table 3

Random Forest Classification*

Ranked Variables From Table 2	Group		
	Control	Alzheimer Disease	Parkinson Disease
All 12	91 (96)	35 (73)	38 (95)
1–8	90 (95)	36 (75)	38 (95)
1–6	87 (92)	32 (67)	35 (88)
1–4	82 (86)	38 (79)	33 (82)

* Data are given as number (percentage) with correct classification by ranked variables in Table 2 compared with expert clinician.