

# Visualization of different characteristics of cerebrospinal fluid with acute encephalopathy and febrile seizures using pattern recognition analysis of $^1\text{H}$ NMR

Takeshi Asano<sup>1</sup>, Keiko Hirakawa<sup>2,3</sup>, Kaoru Koike<sup>4</sup>, Youkichi Ohno<sup>2,3</sup> and Osamu Fujino<sup>1</sup>

**BACKGROUND:** In acute encephalopathy, deterioration of the condition can be rapid, and early intervention is essential to prevent progression of the disease. However, in the acute period, differentiating acute encephalopathy from febrile seizures is difficult. Thus, an early diagnostic marker has been sought to enable early intervention. Proton nuclear magnetic resonance ( $^1\text{H}$  NMR) spectroscopy is used to study the chemical characteristics of biological fluids such as cerebrospinal fluid (CSF). The purpose of this study was to ascertain if pattern recognition of  $^1\text{H}$  NMR spectra could differentiate CSF obtained from patients with acute encephalopathy and febrile seizures.

**METHODS:** CSF was obtained from patients with acute encephalopathy ( $n = 4$ ), complex febrile seizures ( $n = 9$ ), and simple febrile seizures ( $n = 9$ ).

**RESULTS:** NMR spectra of CSF did not visually differ across the three groups. Spectral data were analyzed by partial least squares discriminant analysis and visualized by plotting the partial least squares scores of each sample. The three patient groups clustered separately on the plots.

**CONCLUSION:** In this preliminary study, we were able to visualize different characteristics of CSF obtained from patients with acute encephalopathy and simple and complex febrile seizures using pattern recognition analysis of  $^1\text{H}$  NMR data.

Acute encephalopathy is characterized by the sudden onset of high fever, lethargy, convulsions, and loss of consciousness and has poor prognosis when associated with viral infections, including infections by causative viruses such as the influenza virus, and other etiologies (1–3). Given that encephalopathy leads to rapid deterioration and a poor prognosis, early intervention is essential to prevent progression of the disease and avoid neurologic complications. However, in the acute period, it is difficult to differentiate acute encephalopathy from febrile seizures, especially complex febrile seizures (4–7). Thus, an early diagnostic marker has been sought to enable early intervention. We previously identified a novel

neuropeptide marker of acute encephalopathy in cerebrospinal fluid (CSF) using surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (8).

Proton nuclear magnetic resonance ( $^1\text{H}$  NMR) spectroscopy has been used to study changes in chemical characteristics of biological fluids such as CSF, serum, and urine (9). Pattern recognition with multivariate analysis is a chemometric technique. We hypothesized that pattern recognition of  $^1\text{H}$  NMR spectroscopy data would distinguish the characteristics of CSF obtained from patients with acute encephalopathy, complex febrile seizures, and simple febrile seizures.

## RESULTS

The mean age of patients did not differ among the groups, nor did cell number, total protein, or total sugar content of the CSF (Table 1). NMR spectra of CSF did not visually differ across the three groups. The partial least squares discriminant analysis (PLS-DA) score plot is shown in Figure 1a. Each dot represents an individual sample. Factors 1 and 2 were calculated by PLS-DA algorithm using all 22 sets of  $^1\text{H}$  NMR spectral data. The three patient groups clustered separately on the plot.

Principal components analysis (PCA) is a mathematical procedure that uses orthogonal linear transformation of data to convert possibly correlated variables into uncorrelated primary components. Unlike principal components analysis, which maximizes the variance in a single data set, PLS maximizes the covariance between two data sets by seeking linear combinations of variables from both sets. PLS-DA is a classical PLS regression (with a regression mode), but the response variable is categorical, i.e., indicates the classes/categories of the samples. PLS-DA is often used for classification and discrimination problems (i.e., supervised classification). Factor 1 explains as much of the variability in the data as possible, and factor 2 explains as much of the remaining variability as possible. Only the factors that explain the most variance are retained. The score plot for factors 1 and 2 gives information about the NMR data and summarizes the variation of the data.

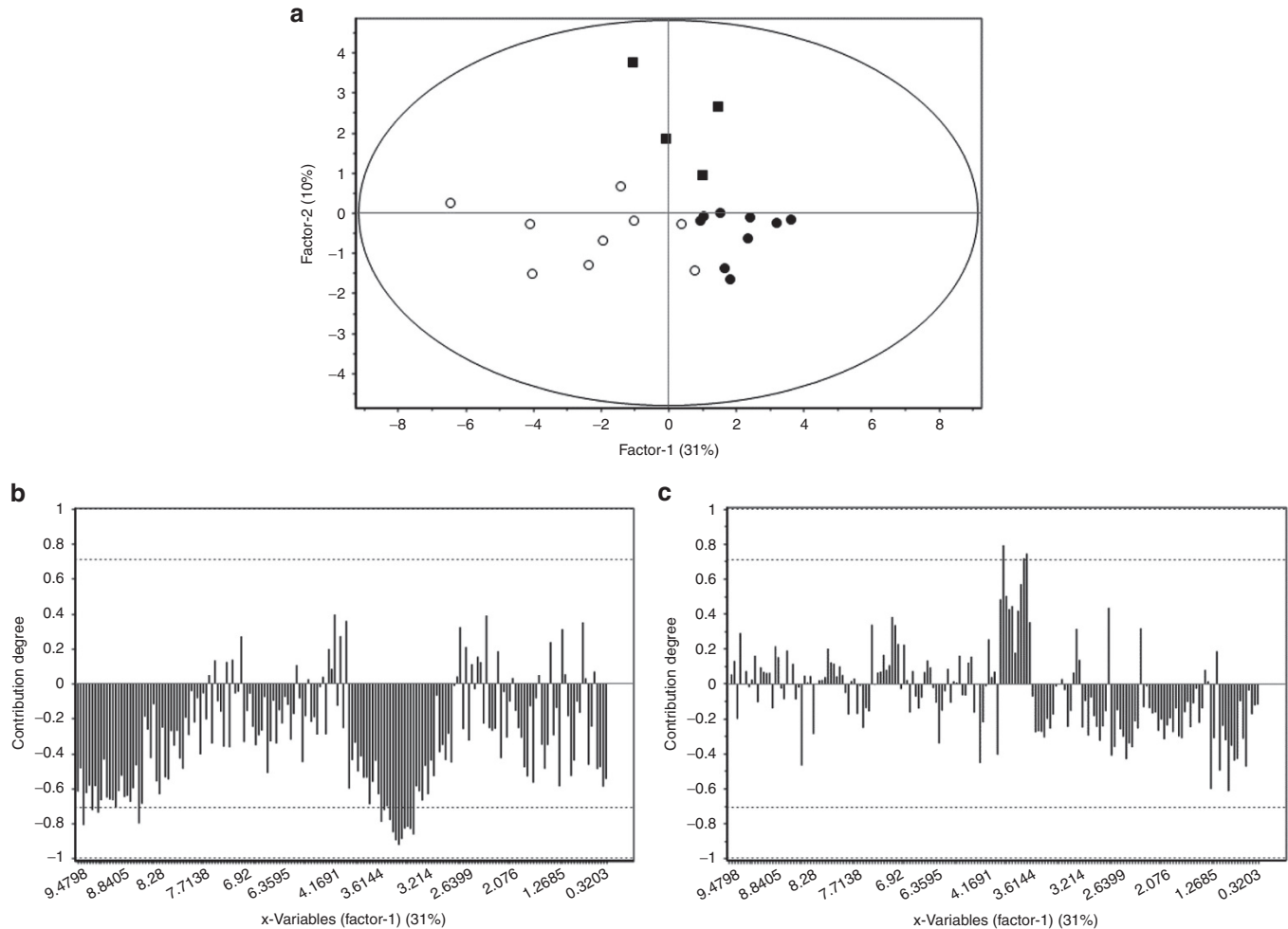
<sup>1</sup>Department of Pediatrics, Nippon Medical School Chiba Hokusoh Hospital, Inzai City, Chiba Prefecture, Japan; <sup>2</sup>Department of Legal Medicine, Nippon Medical School, Bunkyo-ku, Tokyo, Japan; <sup>3</sup>NMR Laboratory, Nippon Medical School, Bunkyo-ku, Tokyo, Japan; <sup>4</sup>Department of Primary Care and Emergency Medicine, Kyoto University Graduate School of Medicine, Sakyo-ku, Kyoto Prefecture, Japan. Correspondence: Takeshi Asano (VFF13540@nifty.ne.jp)

Received 17 February 2014; accepted 17 July 2014; advance online publication 22 October 2014. doi:10.1038/pr.2014.141

**Table 1.** Demographic details of study subjects and CSF characteristics

	Age (years)	Gender (male/female)	Cell number of CSF (/ml)	Total protein in CSF (mg/dl)	Total sugar in CSF (mg/dl)
Complex febrile seizures ( $n = 9$ )	$2.5 \pm 1.7$	7/2	$8.6 \pm 10.2$	$19.0 \pm 5.1$	$76.4 \pm 8.2$
Simple febrile seizures ( $n = 9$ )	$2.9 \pm 2.8$	8/1	$3.9 \pm 3.5$	$22.4 \pm 12.2$	$76.1 \pm 9.8$
Acute encephalopathy ( $n = 4$ )	$5.1 \pm 3.3$	2/2	$3.2 \pm 8.6$	$17.0 \pm 8.8$	$101.1 \pm 29.6$

There is no statistical difference in any of the variables across groups, using a Kruskal–Wallis  $H$ -test with Bonferroni correction. CSF, cerebrospinal fluid.



**Figure 1.** Score and correlation-loading plots on PLS-DA of  $^1\text{H}$  NMR. (a) Score plot based on PLS-DA of  $^1\text{H}$  NMR of CSF from patients with acute encephalopathy, simple febrile seizures, and complex febrile seizures. Each point presents an individual sample. Black squares indicate encephalopathy; open circles indicate simple febrile seizures; filled circles indicate complex febrile seizures. (b,c) Correlation-loading plots for (b) factor 1 and (c) factor 2. The horizontal axis represents x-variables, and the vertical axis represents the degree of contribution. A correlation loading close to outer line indicates that the explained variance of that variable for the displayed components is close to 100%. The inner dashed line corresponds to an explained variance of 50% (hence a correlation of 0.7). CSF, cerebrospinal fluid; NMR, nuclear magnetic resonance; PLS-DA, partial least squares discriminant analysis.

Factor 1 captured 31% of the variance in NMR data and factor 2 captured 10%. Together, these two factors described 41% (31 + 10%) of the variance in NMR data.

Each PLS-DA score plot has a correlation-loading profile that helps assign the spectral regions responsible for the clustering and separation. Corresponding correlation loadings are shown in Figure 1b for factor 1 and in Figure 1c for factor 2, in which the horizontal axis represents the chemical shift and the vertical axis represents the degree of contribution.

The results of the calibration vs. cross-validation analysis are as follows: acute encephalopathy factor 1: the root mean square error for the calibration (RMSEC) = 0.385, the root mean square error for the cross-validation (RMSECV) = 0.407,  $R^2$  value for the explained-variance plot for the calibration data set ( $R^2_c$ ) = 0.004, and  $R^2$  value for the validation data set ( $R^2_{CV}$ ) = not applicable; acute encephalopathy factor 2: RMSEC = 0.216, RMSECV = 0.395,  $R^2_c$  = 0.687, and  $R^2_{CV}$  = 0.034; complex febrile seizures factor 1: RMSEC = 0.355, RMSECV =

0.414,  $R^2_C = 0.478$ , and  $R^2_{CV} = 0.348$ ; complex febrile seizures factor 2: RMSEC = 0.320, RMSECV = 0.423,  $R^2_C = 0.577$ , and  $R^2_{CV} = 0.360$ ; simple febrile seizures factor 1: RMSEC = 0.329, RMSECV = 0.378,  $R^2_C = 0.551$ , and  $R^2_{CV} = 0.431$ ; and simple febrile seizures factor 2: RMSEC = 0.285, RMSECV = 0.373,  $R^2_C = 0.663$ , and  $R^2_{CV} = 0.446$ .

## DISCUSSION

CSF is considered a window into the central nervous system. Spinal dysfunction and brain dysfunction directly affect neuronal metabolism and change the local milieu of the CSF in the spinal column. Previously, we have identified the following four possible biomarker candidates in the CSF for distinguishing acute encephalopathy from febrile seizures using surface-enhanced laser desorption/ionization time-of-flight mass spectrometry (8): (i) a 4.8-kDa proteolytic peptide fragment from neurosecretory protein VGF with a mass-to-charge ratio ( $m/z$ ) of 4,810 (ref. 8), (ii) a double-charged species of transthyretin with  $m/z$  of 6,880 (ref. 10), (iii) ubiquitin with  $m/z$  of 8,564 (ref. 11), and (iv) a truncated form of Cystatin C 1–8 with  $m/z$  of 12,536 (refs. 10,12). In addition, cytokine, chemokine, and protein studies have identified S100 (ref. 13), IL-6 (refs. 14–16), and IL-8 (ref. 17) as possible biomarker candidates for distinguishing acute encephalopathy from febrile seizure. These results strongly suggest that a variety of molecules are related to the differential pathophysiology of acute encephalopathy and febrile seizure, rather than a single substance.

Children with focal and prolonged febrile seizures are more likely to have abnormalities on magnetic resonance images than children with simple febrile seizures, suggesting that the underlying abnormality may predispose the former group of patients to a low seizure threshold (18). Long-term studies have revealed that a history of simple febrile seizures is associated with decreased hippocampal volume and increased  $T_2$  relaxation time more than 15 y after the seizure (19).

Score plots generated using PLS-DA can provide a visual representation of information-rich NMR data by means of dimensionality reduction (20). A score plot shows clustering of biological samples into groups. The closer the samples are in the score plot, the more similar they are with respect to the two components. The plot can be used to interpret differences and similarities among samples. In the current study, the variance was 31% for factor 1 and 10% for factor 2. Factor 1 contributed to separations according to the severity of seizures, and factor 2 contributed to separations between acute encephalopathy and febrile seizures. These combined attributes explained 41% of the variance.

It is helpful to look at the score plot together with the corresponding correlation-loadings plot when determining which variables are responsible for differences between samples. The horizontal axis on the correlation-loadings plot represents original variables, and the vertical axis represents the degree and direction of contribution that influences the separation of scores on the score plot. The higher the loading of a particular variable onto a factor, the more it contributes to that factor. Positive and negative values of loadings have

an opposite power to separate plus or minus direction on the score plot.

In this study, differences in the chemical characteristics of CSF of patients with acute encephalopathy, complex febrile seizures, and simple febrile seizures were visualized using PLS-DA of NMR data. The CSF samples used were obtained at admission when definite clinical diagnosis had not been confirmed. NMR spectroscopy of biofluids such as urine and plasma has been successfully applied to investigate a wide range of diseases and toxic processes (9). Standard one-dimensional NMR spectra typically take only a few minutes to acquire and do not require complex sample preparation. NMR data obtained from CSF contain a variety of chemical information including molecular changes, intermolecular forces, structural changes, and metabolic changes. In this clinical study, the three groups of patients were separated without overlap in the score plot created using NMR data. We validated the discrimination model to estimate how well it would perform on new data; however, the quality of the model evaluated by RMSEC, RMSECV,  $R^2_C$ , and  $R^2_{CV}$  was not sufficiently high.

In conclusion, we used pattern recognition analysis of  $^1\text{H}$  NMR data to visualize different characteristics of CSF obtained from patients with acute encephalopathy and simple and complex febrile seizures. The number of cases used in this study was small, and further research is required.

## METHODS

### Definition of Acute Encephalopathy and Febrile Seizures

Acute encephalopathy is a generic term for acute brain dysfunction usually preceded by infection (1,2). As such, we did not exclude infection-related encephalopathy in this study. Criteria for the diagnosis of encephalopathy were as follows: (i) impaired consciousness, (ii) signs of increased intracranial pressure due to brain edema, such as loss or decreased gray–white differentiation of cortex, hypodensity of thalamus, compressed size of third, fourth, and lateral ventricle, cistern in brain computed tomography image (21), (iii) seizures, (iv) slow activity on electroencephalography lasting more than 24 h after acute onset, and (v) no bacteria or fungi on CSF culture (1,2,8). Simple febrile seizures were defined as primary generalized seizures that lasted for less than 15 min and did not recur within 24 h. Complex febrile seizures were defined as focal seizures that were prolonged ( $\geq 15$  min) and/or recurrent within 24 h (22).

### Subjects

The study was performed on CSF samples obtained from patients with convulsion at the Department of Pediatrics, Nippon Medical School Chiba Hokusoh Hospital from 1 October 2003 to 31 December 2009. CSF samples were obtained from nine children with complex febrile seizures, nine children with simple febrile seizures, and four children with acute encephalopathy. Diagnosis was made by at least two pediatric neurologists based on clinical findings and routine examinations. Indeed, we also do not perform lumbar puncture in all children with simple febrile seizure as well as admission. According to the Guideline for the Neurodiagnostic Evaluation of the Child With a Simple Febrile Seizure by Subcommittee on Febrile Seizures (22), we performed the lumbar puncture in the children with simple febrile seizure who were suspected to have meningeal signs and symptoms (e.g., neck stiffness, and Kernig and/or Brudzinski signs) or in any child whose history or examination suggests the presence of meningitis or intracranial infection. In case of meningitis, we excluded the NMR analysis in this study.

All samples were collected for routine diagnostic tests immediately after hospitalization, and the remaining portion of the samples was used for  $^1\text{H}$  NMR analysis. The Institutional Review Board of the



Nippon Medical School Chiba Hokusoh Hospital approved the collection and investigation of samples, and written informed consent was obtained from the parents or legal guardians of all subjects.

The final diagnosis for four children with acute encephalopathy was as follows: two cases with acute encephalopathy with biphasic seizures and late reduced diffusion, one case with acute encephalitis with refractory, repetitive partial seizures, and one case with acute necrotizing encephalopathy of childhood (1).

### Sample Preparation

All samples were centrifuged at 300g for 5 min to remove cells. Cell-free humoral fractions were divided into aliquots, immediately frozen on dry ice, and stored at  $-80^{\circ}\text{C}$  until further analysis. CSF (200  $\mu\text{l}$ ) was mixed with 340  $\mu\text{l}$  of deuterium oxide ( $\text{D}_2\text{O}$ ; Isotec, St Louis, MO) containing 10 nmol/l of sodium (3-trimethylsilyl) tetradeuterio-propionate-2,2,3,3-d4 (TMSP; MSD Isotopes, Montreal, Canada) and pipetted into 5-mm NMR tubes (Wilmad-LabGlass, Buena, NJ).  $\text{D}_2\text{O}$  provides a deuterium field frequency lock for the NMR spectrometer and TMSP provides an internal chemical shift reference ( $\delta = 0.00$ ).

### Acquisition of $^1\text{H}$ NMR Data

Solution-state  $^1\text{H}$  NMR spectroscopy was performed at a proton resonance frequency of 300 MHz using an ECX-type NMR spectrometer interfaced with a TH5 probe equipped with an automatic 16-position sample changer and Delta NMR processing and control software v. 4.3.2 (all from Jeol Resonance, Tokyo, Japan).  $^1\text{H}$  NMR spectra were acquired automatically at a probe temperature of  $23^{\circ}\text{C}$  using the macro program in the Delta system for automatic measurement. The observation range of the NMR signal was 5,580 Hz. The water resonance was suppressed using a conventional presaturation pulse sequence for the water (HDO) proton signal suppression based on homo-gated irradiation and DANTE pulse sequence (presaturation time = 2 s, DANTE pulse = 8  $\mu\text{s}$ , DANTE interval = 0.1 ms, DANTE loop = 185, DANTE attenuator = 24 dB). Carr-Purcell-Meiboom-Gill spin-echo spectra were measured using a spin-echo loop time of 19.2 ms, a relaxation delay of 2.0 s, and 400 transients.

### Data Processing and Reduction

The resultant spectra were processed using Alice2 (version 5.5; Jeol Resonance). Free induction decays were subjected to an exponential weighing function of 0.2 Hz, Fourier transformed from the time to the frequency domain, and then phased manually, followed by linear baseline correction and referencing to the TMSP singlet at 0.00 ppm.

The signal intensity distribution of  $^1\text{H}$  NMR spectral data was calculated for each sample using a macro program written within Alice2 for Metabolome software (version 1.0; Jeol Resonance) that integrated the spectral data between 0.5 and 9.5 ppm into 225 segments (bins) with 0.04-ppm integral regions. Spectral regions containing resonances of residual water (4.5 to 5.0 ppm) were excluded before integration. Close inspection of the important integral regions for pattern recognition identified some cases where one resonance straddled two or more integral bins, and these integral regions were corrected by manually combining or extending the regions to include one resonance in a single integral region. Other regions that largely consisted of noise were excluded to produce more significant multivariate mapping. Finally, all spectral data were reduced into 174 bins.

### Multivariate and Statistical Analysis

The calculated data obtained from each spectrum were normalized and imported into Unscrambler X software (version 10.3; Camo Software AS, Oslo, Norway) for PLS-DA. The details of the PLS-DA algorithm using Unscrambler X software (23) have been described previously (24). In brief, PLS-DA with cross-validation was performed to investigate the relation between different classes. PLS models use both the  $x$ - and  $y$ -matrices simultaneously to find the latent variables in  $x$  that best predict  $y$ . In other words, PLS-DA was performed to sharpen the separation between groups by rotating components of principal components analysis with the goal of finding maximum separation. In this study, PLS-DA was performed between the spectral profiles of the three groups of patients (acute encephalopathy, simple febrile seizures, and complex febrile seizures). The PLS-DA score plot has a correlation-loading profile that indicates the spectral regions contributing to the sample clustering and separation. The horizontal axis in

the correlation-loading profile represents the spectral regions, and the vertical axis represents the degree of contribution.

Cross-validation was used as a statistical method of evaluating and comparing learning algorithms by dividing data into two segments: one used to learn or train a model and the other used to validate the model. To evaluate the quality of the model, the root mean square error was calculated for the calibration samples (RMSEC) and for the cross-validation samples (RMSECV). To estimate the goodness of fit for future predictions using a defined number of factors, the  $R^2$  value was calculated for the explained-variance plot for the calibration data set ( $R^2_{cv}$ ) and the explained-variance plot for the validation data set ( $R^2_{cv}$ ).

Patient demographic characteristics were compared across groups using a Kruskal-Wallis  $H$ -test with Bonferroni correction, and significant differences were investigated using Mann-Whitney  $U$ -tests.

### ACKNOWLEDGMENTS

The authors specially thank Kaori Okihara, Tomoko Konta, Hideyuki Morikawa, Hideki Muraki, and Tsuyoshi Moriyama for their work on sample preparation, acquisition of  $^1\text{H}$  NMR data, and support of data analysis to complete this research.

### STATEMENT OF FINANCIAL SUPPORT

This work was supported by "The Promotion and Mutual Aid Corporation for Private Schools of Japan" and "The Science Research Promotion Fund." The funders played no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Disclosure: We do not have any financial ties to products in the study.

### REFERENCES

- Mizuguchi M, Yamanouchi H, Ichiyama T, Shiomi M. Acute encephalopathy associated with influenza and other viral infections. *Acta Neurol Scand Suppl* 2007;186:45–56.
- Johnston MV. Acute encephalopathies. In: Behrman RE, Kliegman RM, Jenson HB, eds. *Nelson Textbook of Pediatrics*. 17th edn. Tokyo, Japan: Saunders, 2004:2023–9.
- Sugaya N. Influenza-associated encephalopathy in Japan. *Semin Pediatr Infect Dis* 2002;13:79–84.
- Fishman MA. Febrile seizures. In: McMillan JA, Feigin RD, DeAngelis C, Jones MD Jr, eds. *Oski's Pediatrics Principles & Practice*. 4th edn. Tokyo, Japan: Lippincott Williams & Wilkins, 2006:2297–305.
- Sankar R, Koh S, Wu J, Menkes JH. Paroxysmal disorders. In: Menkes JH, Sarnat HB, Maria BL, eds. *Child Neurology*. 7th edn. Tokyo, Japan: Lippincott Williams & Wilkins, 2006:919–22.
- Millichap JG. The definition of febrile seizures. In: Melson KB, Ellenberg JH, eds. *Febrile Seizures*. New York: Raven Press, 1981:1–3.
- Freeman JM. Febrile seizures: a consensus of their significance, evaluation, and treatment. *Pediatrics* 1980;66:1009.
- Asano T, Koizumi S, Takagi A, et al. Identification of a novel biomarker candidate, a 4.8-kDa peptide fragment from a neurosecretory protein VGF precursor, by proteomic analysis of cerebrospinal fluid from children with acute encephalopathy using SELDI-TOF-MS. *BMC Neurol* 2011;11:101.
- Kim JW, Ryu SH, Kim S, et al. Pattern recognition analysis for hepatotoxicity induced by acetaminophen using plasma and urinary  $^1\text{H}$  NMR-based metabolomics in humans. *Anal Chem* 2013;85:11326–34.
- Ranganathan S, Williams E, Ganchev P, et al. Proteomic profiling of cerebrospinal fluid identifies biomarkers for amyotrophic lateral sclerosis. *J Neurochem* 2005;95:1461–71.
- 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;64:366–70.
- Simonsen AH, McGuire J, Podust VN, et al. A novel panel of cerebrospinal fluid biomarkers for the differential diagnosis of Alzheimer's disease versus normal aging and frontotemporal dementia. *Dement Geriatr Cogn Disord* 2007;24:434–40.
- Shiuhara T, Miyake T, Izumi S, et al. Serum and cerebrospinal fluid S100B, neuron-specific enolase, and total tau protein in acute encephalopathy with biphasic seizures and late reduced diffusion: a diagnostic validity. *Pediatr Int* 2012;54:52–5.

14. Ichiyama T, Morishima T, Isumi H, Matsufuji H, Matsubara T, Furukawa S. Analysis of cytokine levels and NF-kappaB activation in peripheral blood mononuclear cells in influenza virus-associated encephalopathy. *Cytokine* 2004;27:31–7.
15. Ichiyama T, Suenaga N, Kajimoto M, et al. Serum and CSF levels of cytokines in acute encephalopathy following prolonged febrile seizures. *Brain Dev* 2008;30:47–52.
16. Ichiyama T, Ito Y, Kubota M, Yamazaki T, Nakamura K, Furukawa S. Serum and cerebrospinal fluid levels of cytokines in acute encephalopathy associated with human herpesvirus-6 infection. *Brain Dev* 2009;31:731–8.
17. Asano T, Ichiki K, Koizumi S, et al. IL-8 in cerebrospinal fluid from children with acute encephalopathy is higher than in that from children with febrile seizure. *Scand J Immunol* 2010;71:447–51.
18. Reid AY, Galic MA, Teskey GC, Pittman QJ. Febrile seizures: current views and investigations. *Can J Neurol Sci* 2009;36:679–86.
19. Auer T, Barsi P, Bone B, et al. History of simple febrile seizures is associated with hippocampal abnormalities in adults. *Epilepsia* 2008;49:1562–9.
20. Barker M, Rayens W. Partial least squares for discrimination. *J Chemom* 2003;17:166–73.
21. Alper G, Jarjour IT, Reyes JD, Towbin RB, Hirsch WL, Bergman I. Outcome of children with cerebral edema caused by fulminant hepatic failure. *Pediatr Neurol* 1998;18:299–304.
22. Subcommittee on Febrile Seizures. Clinical practice guideline—febrile seizures: guideline for the neurodiagnostic evaluation of the child with a simple febrile seizure. *Pediatrics* 2011;127:389–94.
23. Esbensen KH. *Multivariate Data Analysis in Practice*. 5th edn. Oslo, Norway: CAMO Software, 2010.
24. Wold S, Sjostrom M, Eriksson L. PLS-regression a basic tool of chemometrics. *Chemometr Intell* 2008;58:109–130.