Discriminating poststroke depression from stroke by nuclear magnetic resonance spectroscopy-based metabonomic analysis

Abstract: Poststroke depression (PSD), the most common psychiatric disease that stroke survivors face, is estimated to affect ~30% of poststroke patients. However, there are still no objective methods to diagnose PSD. In this study, to explore the differential metabolites in the urine of PSD subjects and to identify a potential biomarker panel for PSD diagnosis, the nuclear magnetic resonance-based metabonomic method was applied. Ten differential metabolites responsible for discriminating PSD subjects from healthy control (HC) and stroke subjects were found, and five of these metabolites were identified as potential biomarkers (lactate, α-hydroxybutyrate, phenylalanine, formate, and arabinitol). The panel consisting of these five metabolites provided excellent performance in discriminating PSD subjects from HC and stroke subjects, achieving an area under the receiver operating characteristic curve of 0.946 in the training set (43 HC, 45 stroke, and 62 PSD subjects). Moreover, this panel could classify the blinded samples from the test set (31 HC, 33 stroke, and 32 PSD subjects) with an area under the curve of 0.946. These results laid a foundation for the future development of urine-based objective methods for PSD diagnosis and investigation of PSD pathogenesis.

Keywords: poststroke depression, PSD, stroke, nuclear magnetic resonance, NMR, metabonomic

Introduction

Poststroke depression (PSD), the most common psychiatric disease that stroke survivors face, is estimated to affect ~30% of poststroke patients.1,2 Researchers reported that there were 7.5 million stroke survivors in the People’s Republic of China, and the number of PSD patients was ~3 million.1 PSD has deleterious effects on the cognitive and motivation functions of patients and negatively impacts the functional recovery.4 A previous study found that the improvement in depressive symptoms in poststroke patients was associated with better recovery in activities of daily living.5 Thus, an early and accurate diagnosis of PSD is very important.

However, there are still no objective laboratory-based tests to diagnose PSD until now. Currently, PSD diagnosis still relies on the subjective identification of the clinical symptoms, according to the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV) criteria.6 However, this method often results in a considerable error rate in clinical practice, due to the highly clinical heterogeneity of depression. To address this issue, there is an urgent need to develop an objective laboratory-based test for PSD diagnosis.

Metabonomics – the analysis of expression of metabolites in biosamples – could help us better understand the pathophysiological mechanisms of diseases and aid in
the development of diagnostic tools.\(^7\)\(^8\) At present, three major analytical techniques are suited for nontargeted metabonomic mapping, including nuclear magnetic resonance (NMR) spectroscopy, gas chromatography–mass spectroscopy, and liquid chromatography–mass spectroscopy.\(^9\) Many researchers have successfully used metabolomics to identify some potential biomarkers for psychiatry disorders, such as depression, bipolar disorder, and schizophrenia.\(^10\)\(^12\) Meanwhile, Jung et al.\(^13\) reported that there were urinary metabolic abnormalities in stroke patients. These studies indicated that it was possible to use metabolomics to discriminate PSD patients from stroke patients. Therefore, in this study, an NMR spectroscopy combined with multivariate pattern recognition techniques was used to profile the metabolites in urine samples from PSD patients and stroke patients, in order to identify potential biomarkers for PSD diagnosis.

### Patients and methods

#### Study subjects

Totally, 94 PSD and 78 stroke patients were recruited in the Department of Neurology of The First Hospital of Qiqihar City and Central Hospital of Jiamusi City between April 2013 and March 2015. Patients with hemorrhagic or ischemic stroke were included. The stroke diagnosis was performed by a neurologist according to the World Health Organization criteria and confirmed by magnetic resonance imaging.\(^14\) The depression diagnosis was performed by a trained psychologist according to the DSM-IV criteria. The Hamilton Depression Rating Scale was used to assess the severity of depressive symptoms, and the included PSD patients had a score >8.\(^15\)

The PSD and stroke patients who met the following criteria were excluded: 1) preexisting depression prior to stroke; 2) unable to complete the clinical psychological test; 3) concomitant with Parkinson’s disease, inflammatory diseases, psychiatric illness, or severe physical diseases; and 4) illicit drug use or alcohol abuse. Additionally, 74 healthy control (HC) subjects who were without previous lifetime history of DSM-IV Axis I/II and neurological or systemic medical illness were recruited. The clinical characteristics of the included PSD, stroke, and HC subjects are shown in Table 1. Written informed consents were provided by all included subjects. The ethical committee of The First Hospital of Qiqihar City reviewed and approved this study.

#### Sample collection

Urine samples were collected in a sterile cup between 8 am and 10 am after overnight fasting and immediately placed on ice. The samples were centrifuged at 1,500 \(\times\) g for 10 minutes at normal temperature. The resultant supernatant was immediately aliquoted and stored at \(-80^\circ\)C. These procedures were completed within 1 hour of collection.

#### NMR acquisition

The procedure for NMR was performed referring to a previous study.\(^16\) Briefly, there were seven steps: 1) thawing the samples and centrifuging at 1,500 g for 10 minutes; 2) mixing 500 \(\mu\)L sample and 100 \(\mu\)L phosphate buffer and centrifuging at 12,000 rpm for 10 minutes; 3) collecting proton spectra using a Bruker Avance II 600 spectrometer (600.13 MHz \(^1\)H frequency); 4) passing the obtained spectra and baseline referenced to TSP resonance at \(\delta\) 0.0; 5) removing the spectral regions of urea and water resonances (\(\delta\) 4.13–6.8); 6) segmenting the spectra into equal widths using the AMIX package; 6) normalizing the spectral segments in each NMR spectrum to the total sum of the spectral intensity; and 7) importing the normalized integral values into the SIMCA-P 12.0 software.

#### Multivariate pattern recognition

Orthogonal partial least-squares discriminant analysis (OPLS-DA) was used to visualize the discrimination between PSD and stroke/HC subjects. The two parameters (\(R^2\)Y and

---

### Table 1 Demographic and clinical characteristics of included subjects

<table>
<thead>
<tr>
<th>Variable</th>
<th>Training set</th>
<th>Test set</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HC</td>
<td>Stroke</td>
</tr>
<tr>
<td>Sample size</td>
<td>43</td>
<td>45</td>
</tr>
<tr>
<td>Sex (M/F)</td>
<td>24/19</td>
<td>24/21</td>
</tr>
<tr>
<td>Age (year, mean ± SD)</td>
<td>58.9±1.1</td>
<td>60.9±8.7</td>
</tr>
<tr>
<td>Married (Y/N)</td>
<td>33/10</td>
<td>37/8</td>
</tr>
<tr>
<td>Infarction location (left)</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>Hdrs (mean ± SD)</td>
<td>0.39±0.8</td>
<td>0.51±0.9</td>
</tr>
<tr>
<td>BMI (mean ± SD)</td>
<td>21.7±2.4</td>
<td>22.4±2.4</td>
</tr>
</tbody>
</table>

Note: Chi-square analyses for categorical variables and one-way ANOVA test for continuous variables.  
Abbreviations: HC, healthy control; PSD, poststroke depression; M, male; F, female; Y, yes; N, no; Hdrs, Hamilton Depression Rating Scale; BMI, body mass index; ANOVA, analysis of variance; SD, standard deviation.
Q²Y) were used to assess the quality of the built OPLS-DA model. The 199-iteration permutation test was used to rule out the nonrandomness of separation. The coefficient loading plot of the built OPLS-DA model was used to identify the important metabolites contributing to the separation. Based on the number of samples, a correlation coefficient of |r| > 0.250 (equivalent to a P-value of < 0.05) was adopted as a cut-off value. The multivariate logistic regression analysis was used to identify a potential and simplest biomarker panel for PSD diagnosis. The receiver operating characteristic (ROC) curve analysis was used to assess the diagnostic performance of the identified panel. The workflow of this work is described in Figure 1.

Statistical analysis
Mean and standard deviation were used to express the data characterized by a normal distribution. Chi-squared test or one-way analysis of variance (ANOVA) was used when appropriate. If there was a significant difference, Bonferroni or Tamhane’s T2 post hoc test was applied to determine which two groups differed significantly according to the equal variance criterion. For all analyses, P-value < 0.05 was considered to be statistically significant.

Results
Metabonomic analysis
OPLS-DA analysis was used to identify the significantly different metabolites in the urine of PSD subjects compared to stroke and HC subjects. In the training set (43 HC, 45 stroke, and 62 PSD subjects), the score plot of the built OPLS-DA model showed that PSD subjects were distinguishable from stroke and HC subjects with little overlap (Figure 2A). The values of R²Y cum (0.79) and Q²Y (0.65) indicated that this model had a good goodness of fit and predictability. Moreover, the 199-iteration permutation test showed that this model was valid as the corresponding permuted values of R² and Q² were significantly lower than the original values (Figure 2B). In order to validate the discrimination power
of this model, the testing set (31 HC, 33 stroke, and 32 PSD subjects) was used. The T-predicted scatter plot from the built OPLS-DA model showed that only ten of the 64 (15.6%) nondepressed subjects (HC and stroke subjects; Figure 3A) were predicted as PSD subjects and 29 of the 32 PSD subjects (90.6%) were correctly predicted (Figure 3B), yielding a predictive accuracy of 86.4%. These results demonstrated that this model generated by the urinary metabolite profiling might become an empirical diagnostic tool for PSD.

**Differential metabolites**

By analyzing the loading coefficient plot of the built OPLS-DA model, ten differential metabolites ($|r|>0.250$) responsible for the separation of PSD subjects from stroke and HC subjects were identified. Compared to stroke and HC subjects, the levels of the following five metabolites were significantly increased in the urine of PSD subjects: carnosine, arabinitol, α-glucose, phenylalanine, and formate. The levels of the following five metabolites were significantly decreased in the urine of PSD subjects: glutamate, lactate, α-hydroxybutyrate, sarcosine, and acetamide. Meanwhile, the one-way ANOVA was used to validate the metabolites identified by the multivariate statistical analysis. The results showed that the majority of the identified metabolites remained significantly changed (Table 2).

**Important metabolites**

In clinical practice, quantitative measuring of the total ten metabolites was not very convenient and economical. Therefore, the multivariate logistic regression analysis was performed to further decrease the number of metabolites used to diagnose PSD. The forward stepwise regression search algorithm based on Akaike information criterion (AIC) was used. The results demonstrated that the most significant deviations between PSD subjects and HC/stroke subjects could be described by the following five metabolites: lactate, α-hydroxybutyrate, phenylalanine, formate, and arabinitol (Figure 4). These five metabolites could yield a diagnostic accuracy of 83.9% in the training set and a predictive accuracy of 81.3% in the testing set for PSD diagnosis.

**ROC analysis**

To assess the diagnostic performance of the potential biomarker panel consisting of the five identified important metabolites by AIC rule (Figure 5A), ROC analysis was performed. The value of the area under the curve (AUC) was viewed as the assessment index. Generally speaking, the AUC value of 1.0 represented the best diagnostic performance, while the AUC value of 0.5–0.75 represented the weakly diagnostic performance. Here, the AUC value of this panel was 0.946 (95% confidence interval: 0.912–0.979, $P<0.0001$) in the training set (Figure 5B) and 0.946 (95% confidence interval: 0.906–0.987, $P<0.0001$) in the testing set (Figure 5C). These results indicated that the diagnostic performance of this panel was similar to the built OPLS-DA model with all differential metabolites, demonstrating the efficacy of this panel in PSD diagnosis.
Table 2 Important metabolites responsible for discrimination of samples

<table>
<thead>
<tr>
<th>No</th>
<th>Metabolite</th>
<th>Chemical shift/multiplicity</th>
<th>r*</th>
<th>Fold change</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PSD/stroke</td>
<td>PSD/HC</td>
</tr>
<tr>
<td>1</td>
<td>Glutamate</td>
<td>2.35 (m), 3.78 (t)</td>
<td>−0.258</td>
<td>0.66</td>
<td>0.69</td>
</tr>
<tr>
<td>2</td>
<td>α-hydroxybutyrate</td>
<td>1.64 (m), 3.99 (dd)</td>
<td>−0.329</td>
<td>0.53</td>
<td>0.52</td>
</tr>
<tr>
<td>3</td>
<td>Lactate</td>
<td>1.32 (m), 4.10 (m)</td>
<td>−0.587</td>
<td>0.55</td>
<td>0.63</td>
</tr>
<tr>
<td>4</td>
<td>Sarcosine</td>
<td>2.73 (m), 3.60 (m)</td>
<td>−0.306</td>
<td>0.69</td>
<td>0.73</td>
</tr>
<tr>
<td>5</td>
<td>Acetamide</td>
<td>2.05 (m)</td>
<td>−0.300</td>
<td>0.70</td>
<td>0.80</td>
</tr>
<tr>
<td>6</td>
<td>Carnosine</td>
<td>3.19 (m), 7.09 (m)</td>
<td>0.286</td>
<td>1.26</td>
<td>1.29</td>
</tr>
<tr>
<td>7</td>
<td>Arabinitol</td>
<td>3.57 (m), 3.93 (m)</td>
<td>0.262</td>
<td>1.21</td>
<td>1.11</td>
</tr>
<tr>
<td>8</td>
<td>α-glucose</td>
<td>3.42 (t), 3.54 (dd), 3.84 (m), 5.23 (d)</td>
<td>0.292</td>
<td>1.07</td>
<td>1.12</td>
</tr>
<tr>
<td>9</td>
<td>Phenylalanine</td>
<td>3.19 (m), 3.98 (dd), 7.32 (d), 7.36 (m), 7.42 (m)</td>
<td>0.335</td>
<td>2.22</td>
<td>2.02</td>
</tr>
<tr>
<td>10</td>
<td>Formate</td>
<td>8.43 (m), 8.46 (m)</td>
<td>0.378</td>
<td>1.45</td>
<td>1.27</td>
</tr>
</tbody>
</table>

Notes: *Correlation coefficient was obtained from the built OPLS-DA model with a threshold of 0.250. Positive values indicated higher levels in PSD subjects, and negative values indicated lower levels in PSD subjects. >1 values indicated higher levels in PSD subjects, and <1 values indicated lower levels in PSD subjects. P-values were derived from one-way ANOVA.

Abbreviations: PSD, poststroke depression; HC, healthy control; multiplicity: m, multiplet; t, triplet; dd, doublet of doublets; d, doublet; OPLS-DA, orthogonal partial least-squares discriminant analysis; ANOVA, analysis of variance.

Discussion

Depression is the common neuropsychiatric consequence of stroke. PSD has been associated with poor rehabilitation, increased distress, physical disability, and suicidal ideation. Moreover, PSD might affect the acceptability and effectiveness of therapy for patients and increases the length of hospital stay. However, the current PSD diagnosis still remains the subjective identification of depressive symptoms rather than objective laboratory tests, and the underlying molecular mechanism is still unclear. Here, an NMR-based metabonomic method was performed to identify the significantly changed metabolites in the urine of PSD subjects. This method could discriminate PSD subjects from HC and stroke subjects with high accuracy in both the training set and testing set. Finally, a simplified potential biomarker panel consisting of only five metabolites was identified. The accuracy of this panel in discriminating PSD subjects from HC and an AUC of 0.946 in both the training set and testing set. The satisfactory predictability of this panel indicated that it might serve as a “good” classifier for PSD diagnosis.

Here, the levels of ten initially identified metabolites in the urine of PSD subjects were significantly changed, and the OPLS-DA built with these metabolites could provide excellent performance in distinguishing PSD subjects from HC and stroke subjects. This finding demonstrated a robust diagnostic performance of these ten metabolites. These metabolites might aid in the future investigation of the pathogenesis of PSD and the development of an objective diagnostic method for PSD. To make the clinical PSD diagnosis more economical and feasible, a stepwise optimization algorithm based on AIC rule was applied. Finally, a simplified potential biomarker panel consisting of five metabolites was identified. The accuracy of this panel in discriminating PSD subjects from HC and
stroke subjects was still high. Owing to this analytical and rigorous selective strategy, the smaller number of metabolites in the panel could probably have more clinical value than those identified by a previous metabolomic study on PSD diagnosis. Among the five identified potential metabolite biomarkers, the levels of lactate and α-hydroxybutyrate were increased and the levels of phenylalanine, formate, and arabinitol were decreased in the urine of PSD subjects. Although the difference in arabinitol level between PSD subjects and HC/stroke subjects was not statistically significant in the univariate analysis, this metabolite was still identified as an important metabolite and included in the potential metabolite biomarker panel. This was done because the OPLS-DA and binary logical regression analysis showed that the addition of this metabolite could significantly increase the discriminatory and predictive power. These results indicated that the multivariate statistical analysis had an advantage over the univariate analysis in detecting the subtle metabolic differences between different groups. To better understand the underlying molecular mechanisms of PSD, these differential metabolites were analyzed using the online software MetaboAnalyst in terms of metabolic activity in vivo. The results showed that these metabolites were primarily involved in 1) pyruvate metabolism (lactate and formate); 2) propanoate metabolism (lactate and α-hydroxybutyrate); 3) nitrogen metabolism (phenylalanine and formate); and 4) D-glutamine and D-glutamate metabolism (α-glucose). A previous study reported on nonsignificantly changed lactate, α-hydroxybutyrate, and phenylalanine levels in patients with major depressive disorder (MDD), but in this study the lactate and α-hydroxybutyrate levels were significantly decreased and the phenylalanine level was significantly increased. Meanwhile, when using the previously identified biomarker panel for MDD diagnosis to diagnose PSD, its accuracy was inferior to that of the identified panel in this work. These results might indicate that there were different pathogeneses between PSD and MDD and it was not appropriate to use the diagnostic method for MDD to diagnose PSD.

Several limitations should be mentioned in this work: 1) due to the same ethnicity of the included subjects, ethnospecific biases might exist; 2) due to the wide concentration range of human metabolome, this study might miss some of the differential metabolites in the urine of PSD subjects; 3) it was unknown whether the identified panel could be used to discriminate PSD subjects from stroke patients with other psychiatric disorders; and finally, 4) cerebrospinal fluid from PSD subjects should be analyzed to ensure the relationship between these potential biomarkers and disease pathogenesis.

**Conclusion**

This study showed that the NMR-based urinary metabonomics approach might be an effective tool in the clinical diagnosis of PSD. There were ten differential metabolites in the urine of PSD subjects compared to those of HC and stroke subjects. Meanwhile, a potential and an effective biomarker panel consisting of lactate, α-hydroxybutyrate, phenylalanine, formate, and arabinitol was identified. These results laid a foundation for the future development of urine-based objective methods for PSD diagnosis and investigation of PSD pathogenesis.

**Disclosure**

The authors report no conflicts of interest in this work.
References


