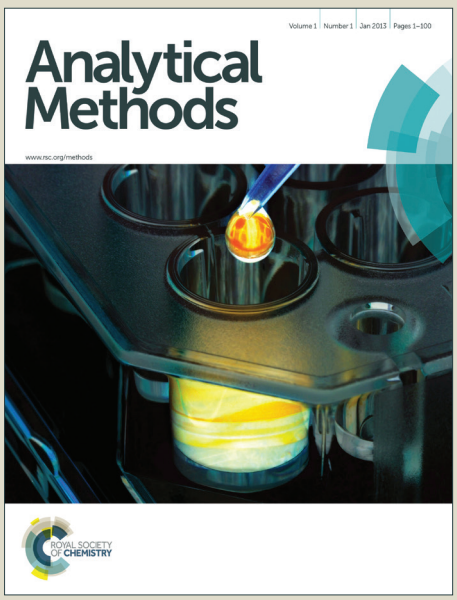
Analytical Methods

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/methods

A novel strategy to identify analytical markers of Cerebralcare Granule for quality assessing by ultra-high performance chromatography and chemometric analysis

Xiao-wen Li[1], Ling Tong[2], Dong-xiang Li[2], Wan-yang Sun[1] , Guo-xiang Sun[1*]

[1]School of Pharmacy, Shenyang Pharmaceutical University, Shenyang 110016, China

[2]Tasly R&D Institute, Tianjin Tasly Group Co., Ltd., Tianjin 300402, China

*Author for correspondence:

Guo-xiang Sun, Ph.D.

School of Pharmacy, Shenyang Pharmaceutical University, 103 Wenhua Road, Shenyang 110016, P. R. China

Tel.: (+86) 24 23986286

E-mail address: sgxwm0416@163.com

Abstract

Selecting appropriate analytical markers of botanical drug is a critical step in quality control. This study presents a strategy to rapidly identify analytical markers of botanical drug via ultra-high performance chromatography coupled with diode array detector (UPLC - PDA) and chemometric methods. Untargeted principle component analysis (PCA) and supervised orthogonal partial least squares discriminate analysis (OPLS-DA) were adopted to select key markers contributed to batch-to-batch variations. Cerebralcare Granule (CG), a commonly used botanical drug for the treatment of cerebrovascular diseases, was employed as an example. Thirty batches of CG samples were analyzed by UPLC to generate dataset of peak areas for PCA and OPLS-DA. Obvious clustering was observed in the OPLS-DA score plot, and candidate markers were recognized from the scatter plots and variable importance in projection values of OPLS-DA. Chemometric analysis extracted eight candidate markers, whose accurate molecular weights were assigned by liquid chromatography combined with quadrupole time-of-flight mass spectrometry. Finally, an UPLC-based method was developed and validated for simultaneous quantification of eight markers in various samples. Using this strategy, rosmarinic acid, gallic acid, chlorogenic acid, peoniflorin, albifiorin, ferulic acid, caffeic acid and aurantio obtusifolin were rapidly identified as the analytical markers of CG. The results of quantitative analysis indicated that the contents of those markers exhibited significantly difference between the qualified and expired samples. In conclusion, the integrated use of UPLC and chemometric analysis provides a reliable approach to the identification of analytical markers for quality control of botanical drug.

Keywords:

Chemometrics; analytical markers; Chromatographic fingerprint; Cerebralcare Granule; multi-component quantification; botanical drug.

Introduction

Analytical markers are frequently identified or quantitated in the quality control of botanical drugs and traditional Chinese medicines (TCMs).[1, 2] According to the definition by the EMEA, an ideal analytical marker for a natural product should not only be characteristic constituent but also be the therapeutic constituent. [3, 4] In many cases, however, it is not clear which components in a botanical product are attributed to the therapeutic outcomes. Currently, one or several components with high contents are selected as chemical markers for developing the quality control approach in most case.[5-7]Since botanical drug and TCMs consist of complex mixtures of phytochemical constituents, a mere of chemical markers will not be sufficient and specific enough for quality control, quality assurance and stability assessment.[8] Hence, a comprehensive approach to identify analytical markers for the quality control of botanical drug and TCMs is required.

Ultra-high performance chromatography (UPLC) has been demonstrated to be a powerful technique for the quality control of TCMs.[9-11]However, due to the complexity of botanical drugs, it is hard to manually process complex multivariate UPLC data to distinguish minor differences among batches.[12] Therefore, chemometric methods, such as principal components analysis (PCA) and orthogonality partial least squares-discriminate analysis (OPLS-DA), etc., provide a good opportunity for mining more useful chemical information from chromatographic data[13-19]; Contribution plots generated by PCA and OPLS-DA have been employed to identify the minor variations chemical markers between different batches products. [20,21] With the help of UPLC and LC–MS, those variations chemical constituents can be identified and quantified by standard components. Hence, it was spontaneous to combine UPLC-based chromatographic analysis and chemometric methods for the rapid identification of chemical markers that represent batch-to-batch variation of botanical drugs.

Cerebralcare Granule (CG) is a widely used botanical drug for the treatment of cerebrovascular diseases including stroke, headache and dizziness (CFDA Approved No.1002004736603642). Clinical observations suggested that CG can improve the cerebral blood flow and blood rheology as well as raise the activity of learning and memory of patient.[22,23] CG is developed from the TCM "Siwutang" and composed of eleven herbs (*Radix angelica sinensis, Rhizoma Chuanxiong, Radix paeoniae alba,Radix rehmanniae preparata, Ramulus uncariaecum uncis, Caulis spatholobi,*

Spica prunellae,Semen cassiae, Concha margaritifera usta, Rhizoma corydalis yanhusuo, Herba asari). According to the official quality standard, the major constituents such as ferulic acid and peoniflorin are selected as markers for the quality control of CG.[24,25] In addition, multi-ingredients quantitative analysis for quality evaluation of CG using various analytical methods such as HPLC-UV, GC-MS has been reported recently, however, the reasons that why those constituents are chosen are not clarified. [26, 27]

The aim of this study is to integrated use of UPLC analysis and chemometric analysis to identify analytical markers that represent the variations in chemical composition of CG between expired and qualified samples. Firstly, we used ultra-high performance liquid chromatography-diode array detection (UPLC-PDA) to generate the chromatograms of CG, whereas a total of 64 common peaks were found. The systematic quantified fingerprint method (SQFM) was used to determine the similarity of samples from each other.[28-30] For further understand the preliminary overview of similarities and differences between failed samples (outliers) and qualified samples, an untargeted multivariate statistical method, i.e. PCA, was performed on the two groups of sample.[31] Peak areas of common peaks were further processed by OPLS-DA analysis to holistically compare the difference among samples. As a result, eight candidate markers were selected by chemometric analysis. The structures of these markers were identified by comparing their MS/MS spectra with those of the reference components or deducing from previous reports. Finally, an UPLC method was developed and validated for simultaneous quantification of eight marker components in various samples.

2. Materials and methods

2.1. Materials

HPLC-grade acetonitrile (ACN) was purchased from Merck (Darmstadt, Germany) and formic acid (FA) was obtained from Fluka (Buchs, Switzerland). Distilled deionized water was purified using a Milli-Q water purification system (Millipore, France). The standard rosmarinic acid (RosA), gallic acid (GA), chlorogenic acid (CGA), peoniflorin (PF), albiflorin (AF), ferulic acid (FA), caffeic acid (CA) and aurantio-obtusifolin (AO) were purchased from the National Institute for the Control of Pharmaceutical and Biological Products (NICPP, Beijing, China). All the standard components have over 98% purity. All the chemicals used, including the solvents, were of analytical grade.

2.2 Apparatus and chromatographic conditions

Chromatographic experiments were performed on a Waters Acuity UPLC system (Waters Corp., Milford, MA, USA) equipped with a 2998 photodiode array detector (PDA) together with a quaternary pump, an auto-sample injector, an on-line degasser and an automatic thermostatic column oven. The MS instrument consisted of a Waters Synapt G2 Q-TOF/MS (Waters Corp., Milford, MA, USA) equipped with an electrospray ionization (ESI) source. Chromatographic separation was performed on Endeavorsil C18 (100 mm×2.0 mm, 1.8 μm particle, Dikma) column. The column temperature was maintained at 30□. 0.1% aqueous formic acid (v/v) (A) and acetonitrile (B) were used as the mobile phase. The gradient elution with the flow rate of 0.25 mL/min was performed as follows: 5% –15% B at 0–9 min, 15–18% B at 9–15 min, 18–20% B at 15–20 min, 20–30% B at 20–26 min, 30–45% B at 26–30 min, 45–90% B at 30–34 min, washing with 90% B at 34-36 min and equilibration with 5% B at 36-40 min. The sample inject volume was 3 μL. The PDA detection was set in the range of 190–400 nm at 1nm/step. Based on the maximum absorption of entire chemical component in the UV spectra of the three-dimensional chromatograms obtained by PDA detection, the multivariate dataset was generated. The wavelengths used to measure the components at the maximum wavelength in their UV spectrum. The change detection wavelength method was used for quantitative analysis: 230 nm for 0–4.55 min, 326 nm for 4.55–8.8 min, 230 nm for 8.8–12 min, 326 nm for 12–24 min, and 280 nm for 24–40 min.

The MS analysis was carried out by the ESI source in both positive and negative ion mode, and full-scan mass range was 100–1200 Da. The source temperature was 100□, and the desolvation gas temperature was 300□. The flow rates of cone and desolvation gas were set at 30 L/h and 600 L/h, respectively. The voltages of capillary, cone and extraction cone in positive ion mode were set at 2.5 kV, 35 V and 5.0 V, respectively, and in negative ion mode, they were set at 2.0 kV, 35 V and 5.0 V, respectively. Leucine enkephalin (m/z 556.2771 in positive ion mode and m/z 554.2615 in negative ion mode) was used as a reference mass. Sodium format was used to set up mass spectrometer calibration in both positive and negative ion mode. MS^E was applied for the MS/MS analysis with the low collision energy of 5 eV and the high collision energy of 25–35 eV.

2.3 Sample preparation

A total of 30 batches of CG (Table.S1) were produced by Tasly Pharmaceutical Co. Ltd (Tianjin, China), including 20 batches of qualified CG sample (S1-S20) and 10 batches expired CG sample (S21-S30). Two grams of CG samples was accurately weighed and ultrasonicated twice with 15 mL of 75% ethanol for 30 min. The extraction was combined and evaporated to 10 mL under vacuum, diluted to volume with adequate ultrapure water in a 25 mL brown volumetric ask and stored at 4□.

Stock standard solutions of the 8 standards, i.e. GA, CGA, CA, FA, PF, AF, RosA and AO were accurately weighted and then dissolved with methanol, respectively. A mixed standard solution was prepared by mixing the individual standard stock solutions and diluting the mixed standard solution with methanol to obtain a series of concentrations for the calibration curves. All the standard solutions were stored at 4□.

The sample extraction and standard solutions were filtered through a 0.22 μm membrane filter before analysis.

2.4 Chemometric analysis

2.4.1. Similarity evaluation

The systematic quantified fingerprint method (SQFM) was used to measure the distribution of chemical constituent among various CG samples. The macro qualitative similarity factor (S_m), the quantitative similarity factor (P_m) and the fingerprint leveling coefficient (α) were calculated to assess the consistency according to the following formulations:

$$S_m = \frac{1}{2} \times (S_F + S'_F) = \frac{1}{2} \left(\frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i^2}} + \frac{\sum_{i=1}^{n} \frac{x_i}{y_i}}{\sqrt{n \sum_{i=1}^{n} \left(\frac{x_i}{y_i} \right)^2}} \right) \quad (1)$$

$$P_m = \frac{1}{2} \times (C + P) = \frac{1}{2} \left(\frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} y_i^2} + \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} y_i} \times S'_F \right) \times 100\% \quad (2)$$

$$\alpha = \left| 1 - \frac{P}{C} \right| = \left| 1 - S'_F \right| \quad (3)$$

From eqn (1) to eqn (3), n represents the peak number, x_i and y_i is the peak area of the ith common constituent existing in the chromatograms of sample and the reference sample. The quality of botanical drug was divided into 8 grades in terms of SQFM criterion, as listed in (Table. S2).

2.4.2 Principal component analysis (PCA)

Principal components analysis (PCA) is an unsupervised method for feature extraction and dimensionality reduction, which is the most commonly used method to reduce data dimensionality and provide an overview of sample distribution.[32,33] There is no training set and input data is classified in an ‘‘unsupervised’’ manner. In the current study, PCA analysis of chromatographic data from CG samples was performed using SIMICA-13.00 software (Umetrics AB, Umea, Sweden). For quality evaluation of CG, the areas of 64 peaks in the chromatography (UV-280nm) of all CG samples were analyzed by PCA. The PCA analysis was implemented by performing singular value decomposition on the data array of the chromatograms, which consisted of a total of n×64 data matrix, n was the number of the samples. Each row represented a CG sample and each column contained the values of 64 characteristic peak areas (Table. S4). The components could be found as the potential analytical markers having the most influence on the discrimination amongst different samples on the PCA loading plots.

2.4.3 OPLS-DA analysis

In order to maximize the separation of the clusters observed and to understand which variables are responsible for the class-separating information, OPLS-DA was performed. We obtained a more sophisticated OPLS-DA model (which separates the systematic variation in X into two parts, one of which is linearly related to Y and the other is orthogonal to Y) with the specific discriminate information among the different groups.[34] And then the V-plot and S-plot were used to find the potential analytical markers between qualified and expired samples.

2.5 Data analysis

All chromatographic data acquired were processed by Empower 3 workstation (Waters Corp., Milford, MA, USA). The peak areas were integrated with “ApexTrack Peak Integration” mode. The retention time (t_R) was ranged from 0 to 40 min with tolerance of 0.01 min and the peak integration parameters were set as automatic. Then, areas of the detected peaks together with their respective retention time were exported as Excel files. The peak areas were normalized (area of

each peak at 280 nm was divided by its sample weight, V/mg) to obtain the datasets. The datasets were auto-scaling prior to modeling. All data with three replicates were reported as mean ± standard deviation.

3 Results and discussion

3.1 Analysis of CG by UPLC

3.1.1 Optimization of the extraction and chromatographic conditions

In order to make a comparison on the chemical constituents, we investigated the extraction solvent (95%, 75%, 50%, 25% (v/v) ethanol and water), column types and determine wavelength (230 nm, 254 nm, 280 nm, 300 nm, 326 nm) for CG samples. The index of the information amount (*Ir*) was adopted to screen the optimal condition for sample extraction and separation, where *Ir* was the index which represents the signal size, signal homogenization and the information amount[35]. (Figure.S1) shows the *Ir* value of items, which suggested that the *Ir* value of 75% MeOH, Column 2 and 280 nm is the highest. Thus, these conditions were adopted in the following study.

The optimized chromatographic condition included: column, Endeavorsil C_{18} column (100 mm×2.0 mm, 1.8 μm particle, Dikma); mobile phase: 5% –15% B at 0–9 min, 15–18% B at 9–15 min, 18–20% B at 15–20 min, 20–30% B at 20–26 min, 30–45% B at 26–30 min, 45–90% B at 30–34 min, washing with 90% B at 34-36 min. and equilibration with 5% B at 36-40 min. The wavelength of 280 nm was selected for the UPLC analysis. Figure.1 shows the 5 detection wavelength chromatogram and the ratio fingerprint of 30 batches of CG samples at 280 nm wavelength.

3.1.2 Method validation

The optimized method was validated in terms of repeatability, precision, and stability tests, and peak 35 was chosen as the reference peak because of its strong signal intensity and high resolution. The precision was investigated by repeated injections of S1 for six times consecutively; the repeatability was determined by analyzing six individual sample solutions; the stability was evaluated by analyzing a single sample solution stored at room temperature for 0, 2, 4, 8, 16, and 24 h. RSDs of the relative migration time and the relative peak area of all the analyses were less than 3.0% for the precision, repeatability, and stability tests. All results indicated that the analysis method was able to meet the requirement of the SQFM analysis.

3.1.3 Similarity evaluation

A total of 30 batches CG samples were detected and their chromatograms were recorded under the optimal conditions. The chromatographic peaks in different samples with the same relative retention time were defined as the common peaks. On the basis of the criterion that the common peaks should be present in all the samples, a total of 64 common peaks were found with moderate peak signals and satisfactory resolution, and also the total peak areas of them were 90% more than all the peak areas in every chromatogram.

To evaluate the batch-to-batch consistency, the chromatograms of 30 samples and the RFP constructed by using the mean value of 20 batches qualified sample chromatograms were analyzed. Similarity parameters S_m, P_m, α, and the final quality grades were calculated based on the RFP. Numerical differences of the quality grade assignment for all the samples are observed in (Table.S2), which manifested that the quality consistency varied notably among various batches. The results of S_m, P_m, α as the variant were performed the hierarchical cluster analysis by using the software SPSS 16.0 (Figure.S2). The dendrogram result showed that the samples could be divided into two clusters: samples S01-S20 were in cluster 1; samples S21-S30 in cluster 2. All of the samples in cluster 1 were all qualified, and the squared Euclidean distance was ≤ 4. In cluster 2, all samples were far from cluster 1. The distance between samples increased gradually, just like the change in CG. This indicated that the SQFM could be used to discriminate the expired samples from qualified samples.

3.2 Chemometric analysis

PCA was performed allowing visualization of holistic distribution of the CG products and further evaluation of the quality consistency for those samples. A two-component PCA model was obtained which cumulatively accounted for 66.8% of the variation; the total variance explained for PC1 is 43.1% and that for PC2 is 23.7%. The score plot of the first and second principal components visually showed the obvious differentiation between CGs from different batches (Figure.2). Cluster1 (Gr.1), which were samples S1-S20, and Cluster 2 (Gr.2) represented samples S21-S30, except S28, were clearly clustered. As shown in Figure.2B, The preferential distribution of potential markers were G3, G32, G14, G15, G4, G10, G4, G21, G46, G55, which accounted for the most variation in the dataset and exhibited the greatest correlation to the CG samples. However,

detailed information regarding components contributing to the data differentiation between expired and quantified samples still needs further validation.

The OPLS method was a recent modification of the PLS method, and was a supervised OPLS-DA statistical model, which was carried out to focus on the discrimination of Gr.1 and Gr.2. All OPLS-DA models exhibited R^2 and Q^2 values > 0.8. After removing S28, an outlier with very significant variance, the OPLS-DA model was performed, and a predictive ability $Q^2(Y)$ of 86.6% for Gr. 1 and Gr. 2 was obtained. As shown in Figure.3A, the original two groups Gr. 1 and Gr. 2 were clearly distinguished. To determine which components contributed highly to samples clustering OPLS-DA was used. A scatter plot (S-plot) and variable importance in projection values (VIP) were used to select the significant components that were influence the consistency among differnt samples.(Figure.3B) Potential analytical markers were selected via S-plots by applying a cutoff for correlation (p [corr] at > |0.5| and covariance (w^*[1] at > |0.2|). The VIP plot, displaying the VIP values of all of the variables, the peak whose VIP was larger than 1.50 was selected as the analytical marker. As shown in Figure.3C. There are eight peaks (3, 4, 10, 14, 15, 21, 32, 55) with confidence intervals was individually as analytical markers for discriminate the differentiation between expired and quantified samples.

3.3 Identification of characteristic peaks

Due to the results of OPLS-DA analysis, eight peaks were selected from the 64 common peaks. For holistic quality control of CG, the next step was to detailed clarification of those constituents based on the high-accuracy UPLC-Q-TOF/MS analysis. In the ESI-MS experiment, the molecular weight of each peak and some fragments could be obtained. In our study, both ESI in positive and negative mode were used. Most of the m/z data were [M+H]+ or [M+NH4]+ in positive mode as well as [M−H]- in negative mode. By comparing the retention time, UV and characteristic in ESI-MS spectrum of each peak with the literatures,[36] eight peaks which selected from the results of OPLS-DA were unambiguously identified by standard substances. Typical MS total ion current (TIC) chromatogram traces with numbered peaks are illustrated in Figure.4. Information such as retention times (t_R), MS fragmentation behaviors and distribution are summarized in Table.1.

3.4 Simultaneous quantitative analysis of the marker components

3.4.1. Selection of detection wavelength

To verify the applicability of the selected marker candidates, we analyzed 30 different batches of CG samples. A sensitive and reliable UPLC-based analytical method was established to quantify the eight marker components (GA, CGA, CA, FA, PF, AF, RosA and AO). Because the maximum wavelength of the eight analytes are different, three detection wavelengths of 230nm, 280nm and 326 nm were selected (Figure.5).

3.4.2 Method validation of quantitative analysis.

The linearity of the UPLC method was evaluated at six concentration levels of the marker components under the optimal separation conditions. Table.2 has list linear equation and its correlation coefficient, liner range, limit of detection (LOD) and limit of quantification (LOQ) of each component. All calibration curves were of good linearity with high correlation coefficient ($R^2 \geq 0.99$) over the tested range. The RSD for repeatability of the proposed method was 2.31%. Table.3 shows the results of precision, repeatability and accuracy. Both RSD values of intra- and inter- day precisions did not exceed 2.87%. The average recoveries of each compound were between 95.92% and 105.46%. The validation results demonstrated that the method was acceptable for subsequent analysis of all the samples.

3.4.3 Quantitative analysis of analytical markers in CG samples

The validated UPLC-PDA method was successfully applied to simultaneous determination of the eight marker components (GA, CGA, CA, FA, PF, AF, RosA and AO) in the 30 CG samples. The analysis was conducted in triplicate and the results are shown in Table.4. It is noteworthy that the content of those components were significantly different between expired samples and qualified samples，for example, RosA and GA were present higher mean value in the qualified samples (470 and 234 μg/g) than the expired samples (444 and 156 μg/g). In the results of the expired samples, the contents of these eight components were significantly reduced. This indicated the severe degradation of the eight analytical markers in the expired samples resulting in unstable quality and efficacy. Therefore, the quality of commercial products of CG can be well evaluated and interpreted by those analytical markers.

Conclusion

In this research, UPLC coupled with chemometric analysis provides a novel strategy for rapid identification of analytical markers to assessing the quality of CG. The UPLC-PDA offers a

powerful tool for quantifying and separating the individual components and creates characteristic chromatographic profile. And the quality consistencies of 30 CG samples from the same manufacturer were well differentiated based on SQFM method. PCA and OPLS-DA analysis were used to discriminate samples with tremendous quality variations as outliers, which were beneficial for the exploration of characteristic analytical markers for quality control. As a result, eight marker components were selected as quality control markers in CG, which were simultaneously determined in 30 CG samples. These markers were significantly reduced at the expired samples, which suggested that the selected analytical markers are practical and reliable for quality evaluation. All the results indicate that the present approach is rapid and reliable for analytical marker selection and can be an appropriate means of quality control of TCMs with botanical drugs.

Acknowledgments

This work is supported by National Key Special Project of Science and Technology for Innovation Drug of China (Grant No. 2013ZX09402202).

Conflicts of Interest

The authors declare no conflict of interest.

Reference

1. J. Song, S. Li, Y. Zhou, C. Qiao, S. Chen, H. Xu, *J. Pharm. Biomed. Anal.,* 2010, 53, 279–286.

2. X. Liu, Z. Wu, K. Yang, H. Ding and Y. Wu, *J Pharm Biomed Anal*, 2013, 76, 70-74.

3. The European Medicines Agency: Reflection paper on markers used for quantitative and qualitative analysis of herbal medicinal products and traditional herbal medicinal products. [http://www.emea.europa.eu/pdfs/human/hmpc/25362907en.pdf].

4 . S. Li, Q. Han, C. Qiao, J. Song, C. Lung Cheng and H. Xu, *Chin Med*, 2008, 3, 7.

5. Y. Yang, H. J. Wang, J. Yang, A. H. Brantner, A. D. Lower-Nedza, N. Si, J. F. Song, B. Bai, H. Y. Zhao and B. L. Bian, *J Chromatogr A*, 2013, 1321, 88-99.

6. L. W. Qi, J. Cao, P. Li and Y. X. Wang, *J Pharm Biomed Anal*, 2009, 49, 502-507.

7. X. Shi, K. Zhang, N. Xue, L. Su, G. Ma, J. Qi, Y. Wu, Q. Wang and Q. Shi, *Food Chem*, 2013, 141, 4019-4025.

8. P. S. Xie and A. Y. Leung, *J Chromatogr A*, 2009, 1216, 1933-1940.

9. G. Jin, X. Xue, F. Zhang, Y. Jin, X. Liang, *Talanta*, 2009, 78, 278-283.

10. I.Klimczak, A.Gliszczynska-Swiglo, *Food Chem*, 2015, 175, 100-105.

11. Dadáková. Eva, Křížek.Martin, Pelikánová, Tamara, *Food Chem*, 2009,116,365-370.

12. S. A. Wren, *J Pharm Biomed Anal*, 2005,38,337-343.

13. N. Hakimzadeh, H. Parastar, M. Fattahi, *J Chromatogr A*, 2014, 1326, 63-72.

14. H. Parastar, M. Jalali-Heravi, H. Sereshti, A. Mani-Varnosfaderani, *J Chromatogr A*, 2012, 1251, 176-187.

15.J. M. Amigo, M.J. Popielarz, R. M. Callejón, M. Morales, A. M. Troncoso, M. A. Petersen, T. Toldam-Andersen, *J Chromatogr A*, 2010, 1217, 4422-4429.

16. G. X. SUN, Y. Wu, Z. B. Liu, Y. F. Li and Y. Guo, *Anal.Methods*, 2014,6,838-849.

17. M. A. Farag, A. Porzel and L. A. Wessjohann, *Phytochemistry*, 2012, 76, 60-72

18 S. Yudthavorasit, K. Wongravee and N. Leepipatpiboon, *Food Chem*, 2014, 158, 101-111.

19. D. M. Kulakowski, S. B. Wu, M. J. Balick and E. J. Kennelly, *J Chromatogr A*, 2014, 1364, 74-82.

20. Z. Q. Shi, D. F. Song, R. Q. Li, H. Yang, L. W. Qi, G. Z. Xin, D. Q. Wang, H. P. Song, J. Chen, H. Hao and P. Li, *J Chromatogr A*, 2014, 1345, 78-85.

21. H. Zhu, C. Wang, Y. Qi, F. Song, Z. Liu and S. Liu, *Talanta*, 2013, 103, 56-65.

22.F. Wang, Q. Hu, H. C. Chen, X. S. Xu, C. M. Zhou, Y. F. Zhao, B. H. Hu, X. Chang and J. Y. Han, *Microcirculation*, 2012, 19, 260-272.

23. X. S. Xu, Z. Z. Ma, F. Wang, B. H. Hu, C. S. Wang, Y. Y. Liu, X. R. Zhao, L. H. An, X. Chang, F. L. Liao, J. Y. Fan, H. Niimi and J. Y. Han, *Shock*, 2009, 32, 201-209.

24. P. Huang, C. M. Zhou, H. Qin, Y. Y. Liu, B. H. Hu, X. Chang, X. R. Zhao, X. S. Xu, Q. Li, X. H. Wei, X. W. Mao, C. S. Wang, J. Y. Fan and J. Y. Han, *Exp Neurol*, 2012, 237, 453-463.

25.G. H. Chen, L. Tong, T. J. Hang, Z. W. Wang and Y. F. Wang., *Chin J Pharm Anal*, 2013, 33, 124-127.

26. W. B. Li, J. P. Han, Q. Ni. And J. Gao., *Chin J Pharm Anal*, 2011, 31, 1385-1388.

27. X.P. Chen, X. H. E, Z. T. Xia, X.Y. Zhang, L. L. Zhang, S. P. Zhou, Y. H. Zhu and C. X. Lu., *Chinese Traditional Patent Medicine*, 2013, 35, 1921-1925.

28. W. B. Li, J. P. Han, Q. Ni. And J. Gao., *Chinese Traditional and Herbal Drugs*, 2011, 42, 1321-1325.

29. G. X. SUN, X. Z. ZHI, C.L. ZHANG and H. Y. DONG., *Central South Pharmacy*, 2007, 6, 2.

30. Y. Wang, G. Sun, Z. Liu, Y. Liu, Y. Gao, J. Zhang, Z. Ji and X. Chen, *J Sep Sci*, 2014, 37, 3571-3578.

31. J. Kumirska, N. Migowska, M. Caban, A. Plenis, P. Stepnowski, Journal of Chemometrics, 2011, 25, 636-643.

32. J. Kumirska, A. Plenis, P. Lukaszewicz, M. Caban, N. Migowska, A. Białk-Bielińska, M. Czerwicka, P. Stepnowski, J Chromatogr A, 2013, 1296, 164-178.

33. G. A. Theodoridis, H. G. Gika, E. J.Want, I. D. Wilson, Anal Chim Acta, 2012, 711, 7-16.

34. U. Jumhawan, S. P. Putri, Yusianto, E. Marwani, T. Bamba and E. Fukusaki, *J Agric Food Chem*, 2013, 61, 7994-8001.

35. G. X. SUN and J. X. ZHANG., *Chinese Journal of Analytical Chemistry*, 2009, 8, 25.

36. X. P. Chen, L. Tong, Y. Chu, X. Y. Wang, L. L. Zhang, X. H. Ma, S. P. Zhou and C. X. Liu, Asian Journal of Chemistry, 2013,25, 7840-7842.

Figure captions

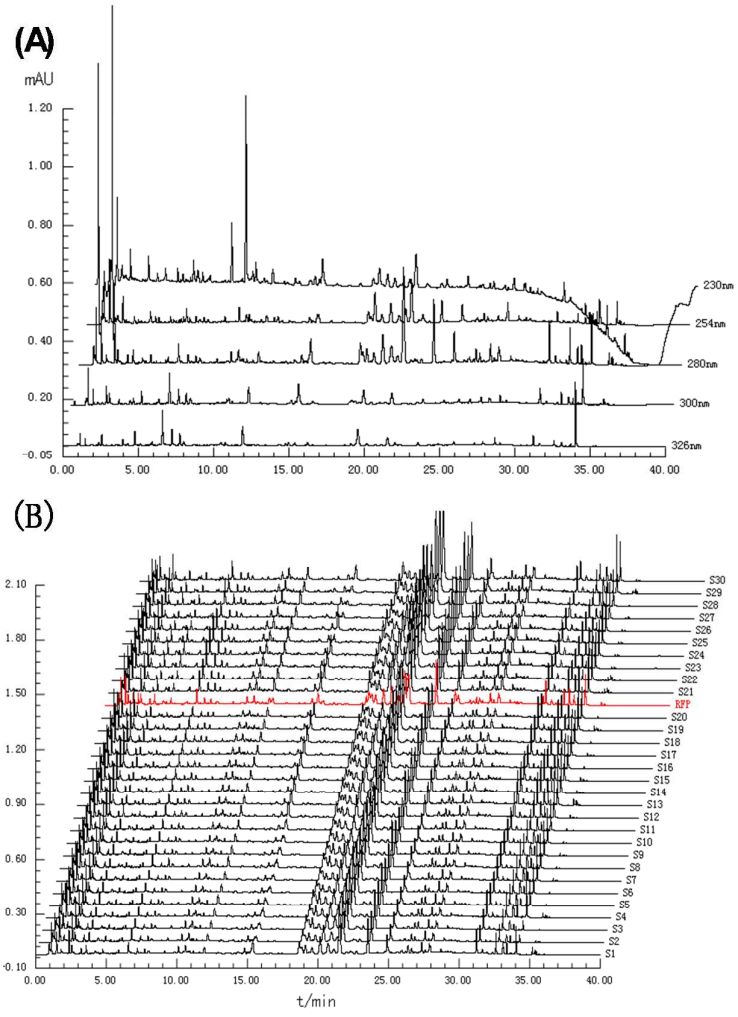


Figure.1 UPLC-PDA chromatography at different UV detector wavelength (A) and the chromatographic fingerprints for 30 batches CG samples (B). The chromatogram of (A) five UV detector wavelength: 230nm, 254nm, 280nm, 300nm, 326nm and (B) standardized chromatographic fingerprint of 30 batches of Cerebralcare Granule samples were obtained by UPLC-PAD method at 280 nm; The RFP represents the reference fingerprint.

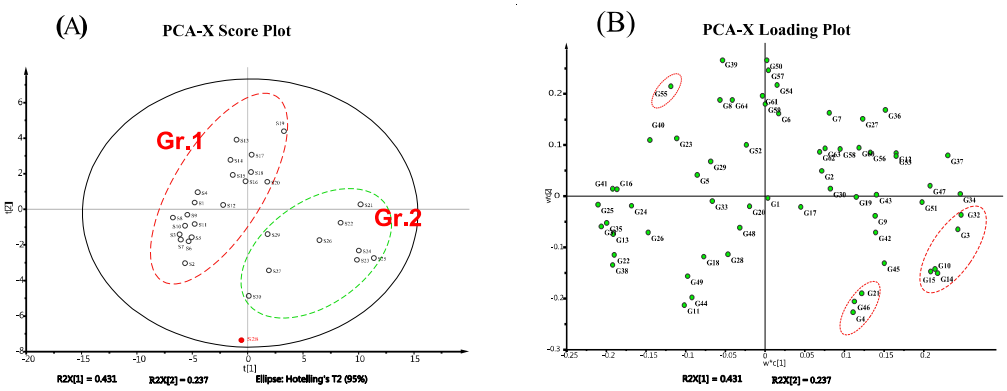


Figure.2 (A)The PCA Score Plot given by SIMCA(13.0 version). *x*-Axis represents the first principal component which explains 43.1% of the total variance. *y*-Axis represents the second principal component which explains 23.7% of the total variance. (B) 2D Loading Plot of the 30 batches CG samples.

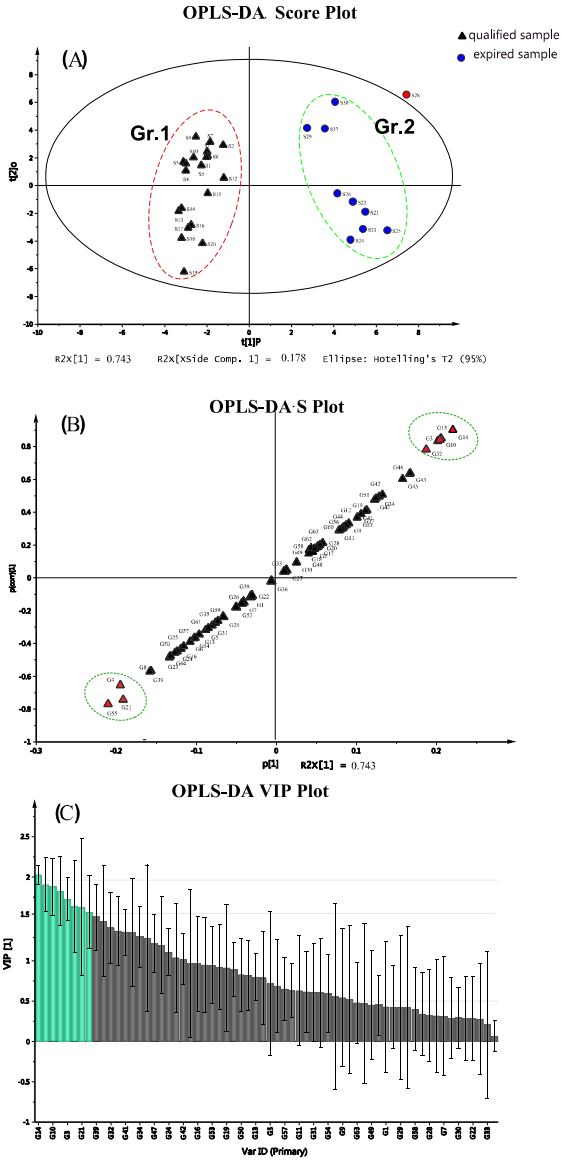


Figure.3 (A) The OPLS-DA score plots of Gr. 1 and Gr. 2 given by SIMCA(13.0 version), drawn with Hoteling's 95% confidence ellipse; (B) S-plot from OPLS-DA. Covariance and correlation plotted on the *x*- and *y*- axis, respectively. (C) VIP plot from OPLS-DA, compounds 3, 4, 10, 14, 15, 21, 32, 55 with larger VIP values were selected as analytical markers.

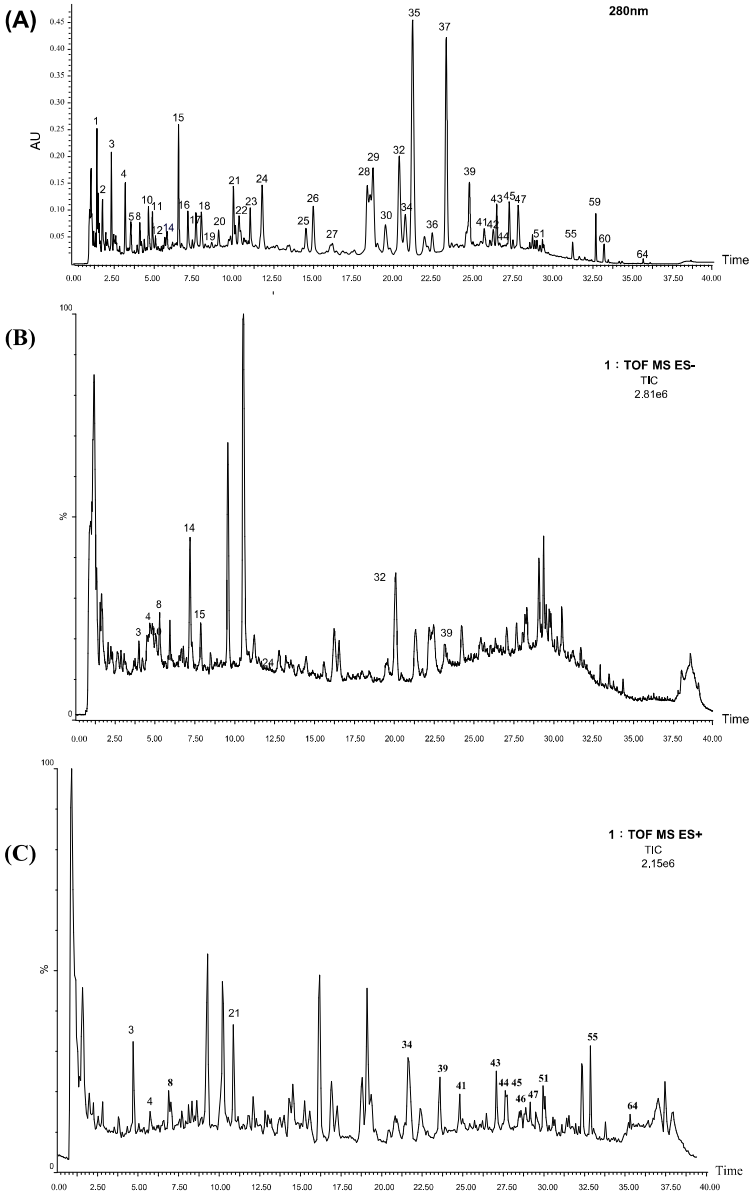


Figure.4 (A) The UPLC–UV chromatography the 64common peaks at 280 nm; (B) the total ion chromatography (TIC) of CG in negative ion mode. (C) and the TIC of CG in positive ion mode.

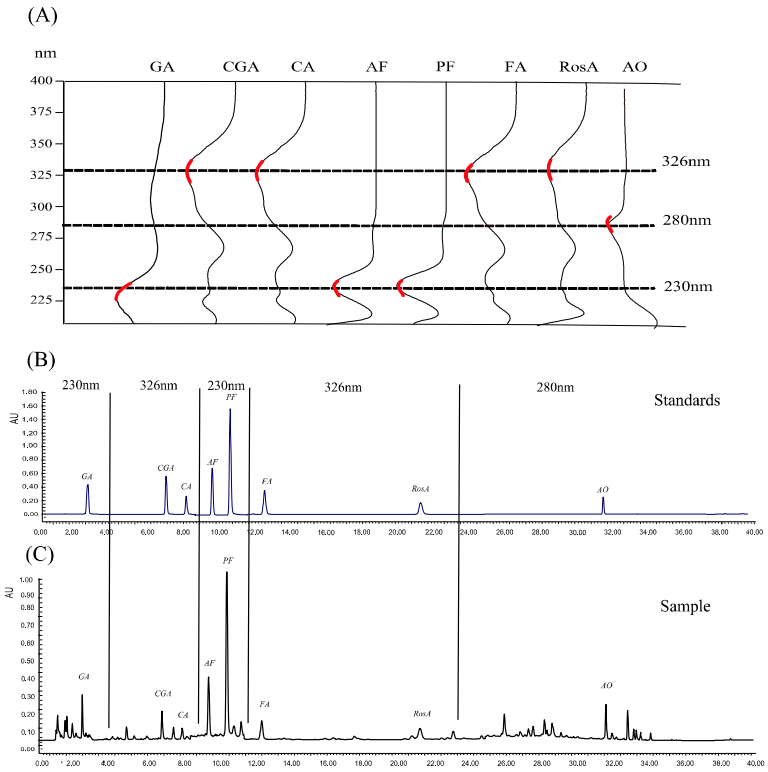


Figure.5 (A) UV spectrum of standards. The chromatogram of (A) mixed standard components : gallic acid (GA), chlorogenic acid (CGA), caffeic acid (CA), albiflorin (AF), peoniflorin (PF), ferulic acid (FA), rosmarinic acid (RosA), and aurantio obtusifolin (AO); (B) UPLC chromatogram of eight mixed reference standards; (C) UPLC chromatogram of Cerebralcare Granule sample solution.

Table

Table.1 Tentative identification of marker components from Cerebralcare Granule.

Peaks NO.	t_R (min)	Measured mass (m/z)	Calculated mass (m/z)	Error (ppm)	Formula	ESI-MS^E (m/z)	Identificationa[a]
3	2.69	169.0133	169.0137	2.4	$C_7H_6O_5$	125.0243 [M-H-CO_2]-,	Gallic acid
4	4.71	353.0863	353.0873	2.8	$C_{16}H_{18}O_9$	191.0507[M-H-caffeoyl]-,179.0343[M-H]-,135.0439[M-H-CO_2]-	Protocatechuic acid
10	8.48	179.0337	179.0344	3.9	$C_9H_8O_4$	135.3018[M-H-CO_2]-	Chlorogenic acid
14	9.56	479.1549	479.1553	0.8	$C_{23}H_{28}O_{11}$	525.1732[M+HCOO]-,357.0928[M-H-BA]-,283.0750[BA+Glc-H]-	Albiflorin
15	10.5	479.1549	479.1553	0.8	$C_{23}H_{28}O_{11}$	449.1436[M-H-HCOH]-,327.1061[M-H-HCOH-BA]-	Paeoniflorin
21	12.17	193.0501	193.0501	0.0	$C_{10}H_{10}O_4$		Ferulic acid
32	21.33	359.0758	359.0767	2.5	$C_{18}H_{16}O_8$	179.9526[M-H-$C_9H_8O_4$]-	Rosmarinic acid
55	32.06	331.0819	331.0818	-0.3	$C_{17}H_{14}O_7$	317.0799[M+H-CH_3]+	Aurantio-obtusin

a Structures confirmed by comparison with reference standards.

Table.2 Calibration curves, LODs and LOQs and UV absorption values of the components.

Component	Standard curve *y=ax+b*	R^2	Linear range(μg/mL)	Wavelength (nm)	LOD(ng/mL)	LOQ(ng/mL)
GA	$y = 2.3\times10^6x + 1988.1$	0.9999	38.44-1230	230	5.00	16.65
CGA	$y = 3.6\times10^6x + 2118.9$	0.9997	6.875-220	326	0.89	2.98
CA	$y = 1.5\times10^6x - 4969.3$	0.9998	20.3-650	326	2.64	8.80
AF	$y = 1.5\times10^6x + 16060$	0.9972	125.2-2040	230	11.82	39.40
PF	$y = 1.0\times10^6x + 38609$	0.9997	187.5-6000	230	24.37	81.23
FA	$y = 4.4\times10^6x + 44096$	0.9943	6.88-220	326	0.89	2.98
RosA	$y = 1.9\times10^6x - 4969.3$	0.9990	62.8-2010	326	8.16	27.21
AO	$y = 3.5\times10^6x + 6986.6$	0.9997	20.6-330	280	1.34	4.47

y, the peak area, *x*, the concentration of each reference compound (μg/mL); R^2, the square of correlation coefficient of regression equations; LOD, limit of detection (S/N=3). LOQ, limit of quantitation (S/N=10). rosmarinic acid (RosA), Gallic acid (GA), Chlorogenic acid (CGA), peoniflorin (PF), Albiflorin (AF), Ferulic acid (FA), Caffeic acid (CA) and aurantio-obtusifolin (AO).

Table.3 Precision and recovery for eight analytical marker in Cerebralcare Granule samples by UPLC-PAD (280 nm).

Component	Intra-day Precision (n=6, %)	Inter-day Precision (n=6, %)	Repeatablity (RSD%,n=6)	Recovery[a](n=3)			
				Original/mg	Spiked/mg	Found/mg	Recovery (%)

					0.37	0.83	100.97±2.12
GA	0.72	1.74	1.96	0.46	0.45	0.93	104.31±1.96
					0.56	1.02	99.58±2.31
					0.27	0.60	99.05±1.15
CGA	2.08	1.39	0.87	0.33	0.30	0.62	97.91±0.56
					0.38	0.72	100.50±1.76
					0.22	0.50	103.71±4.15
CA	1.65	0.78	1.04	0.27	0.29	0.57	101.55±1.27
					0.34	0.62	101.49±3.07
					0.33	0.74	101.91±1.79
AF	1.69	0.81	1.07	0.41	0.41	0.81	98.77±1.63
					0.48	0.91	104.27±3.24
					1.65	3.62	96.32±2.01
PF	0.25	0.70	1.56	2.03	2.01	4.02	98.84±1.68
					2.44	4.49	100.89±3.06
					0.19	0.42	95.92±1.54
FA	0.98	0.62	1.49	0.24	0.25	0.49	97.21±1.63
					0.31	0.55	99.50±2.29
					0.80	1.82	105.46±2.83
RosA	1.76	2.89	2.05	0.98	0.97	1.98	103.50±1.72
					1.14	2.11	99.13±1.07
					0.20	0.45	99.01±2.40
AO	1.89	0.77	2.31	0.25	0.25	0.49	96.23±3.85
					0.30	0.55	100.80±1.85

a Recovery (%)=[1–(concentration found–concentration spiked)/concentration spiked]×100%.

Table.4 The results of quantitative analysis of the marker components in 30 batches sample (μg/g).

Sample[a]	GA	CGA	CA	AF	FP	FA	RosA	AO
S1	271	290	268	1059	246	187	490	126
S2	222	263	637	1153	248	85	651	156
S3	226	383	207	1023	215	198	579	177
S4	275	402	271	1035	220	200	592	231
S5	242	388	241	927	215	233	567	161
S6	295	298	212	1134	250	210	402	169
S7	244	233	219	1060	228	160	294	178
S8	292	360	263	1100	234	223	554	198
S9	284	265	222	1122	247	170	454	205
S10	282	276	240	1138	247	188	371	194
S11	237	120	250	934	204	88	678	183
S12	215	170	224	911	211	151	457	176
S13	296	140	187	972	229	125	411	142
S14	318	164	236	927	202	134	444	154
S15	224	164	285	1177	199	152	404	203

S16	306	158	235	894	184	132	425	113
S17	309	169	264	901	186	133	359	139
S18	300	155	178	873	226	121	417	106
S19	299	151	207	869	199	125	454	141
S20	219	141	166	799	202	140	404	171
Mean1	**268**	**234**	**251**	**1000**	**220**	**158**	**470**	**166**
S21	230	165	137	1015	215	120	399	131
S22	222	192	164	1088	208	119	445	181
S23	260	131	164	1118	215	119	333	147
S24	230	165	226	854	215	142	554	102
S25	207	142	141	839	215	142	457	111
S26	226	128	142	703	208	122	387	97
S27	233	139	173	806	213	129	405	89
S28	256	171	328	1164	206	165	467	92
S29	250	164	219	849	203	143	494	126
S30	229	157	232	851	186	151	501	77
Mean 2	**234**	**156**	**193**	**929**	**209**	**135**	**444**	**115**

a The data are presented at the average of duplicates.