

Jiang Cheng (Orcid ID: 0000-0001-8454-749X)

Prediction of Allergic Reaction Risk of Shenmai Injection Based on Real-World Evidence Coupled with UPLC-Q-TOF-MS

Running title: Prediction of Allergic Reaction Risk of Shenmai Injection

**Cheng Jiang^{1,2,a}, Lisha Zhao^{2,a}, Chuanwei Xin¹, Yu Dong², Jie Shen³, Zhongni Xia¹,
Dan Shou^{2,*} and Gonghua Li^{1,**}**

¹ Department of Pharmacy, Tongde Hospital of Zhejiang Province, Hangzhou 310012, China

² Department of Medicine, Zhejiang Academy of Traditional Chinese Medicine, Hangzhou 310007, China

³ Department of Pharmacy, Zhejiang Province Hospital of Traditional Chinese Medicine, Hangzhou 310006, China

* Correspondence:

Dan Shou, Department of Medicine, Zhejiang Academy of Traditional Chinese Medicine, Hangzhou 310007, China. Email: shoudanok@163.com

** Correspondence:

Gonghua Li, Department of Pharmacy, Tongde Hospital of Zhejiang Province, Hangzhou 310012, China. Email: ligonghua88@163.com

^a These authors contributed equally to this work.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/bmc.5255

Abstract

The allergic reaction (AR) of Chinese herbal injection has become one of the most noticeable focuses of public health in China. However, it still has been remained considerable controversial whether low-molecular-weight components in Chinese herbal injection have potential sensitization. In this study, the relationship between AR and low-molecular-weight component profile of Shenmai injection was explored by an interdisciplinary technology integrating real-world evidence and ultra performance liquid chromatography-quadrupole time-of-flight mass spectroscopy (UPLC-Q-TOF-MS). The AR information of hospitalized patients was obtained by comprehensive analyzing real-world evidence from January 2015 to June 2019 at two Chinese hospitals. The UPLC-Q-TOF-MS was exploited to systematically investigate the low-molecular-weight component profile with 50-1500 m/z mass range and 3,725 MS1 peaks were detected. The optimized partial least squares discriminant analysis model was established to map the influence of low-molecular-weight components on AR. The results of this study showed that high levels of organic acids might be potential risk factor for inducing AR administered intravenously. By using the method, Shenmai injection with high AR risk could be recognized precisely with 100% accuracy before clinical use.

KEYWORDS

Allergic reaction; Shenmai injection; Real-world evidence; Organic acid; UPLC-Q-TOF-MS

Abbreviations: AR, allergic reaction; CHI, Chinese herbal injection; HPLC, high performance liquid chromatography; UPLC-Q-TOF-MS, ultra performance liquid chromatography quadrupole time-of-flight mass spectroscopy; PLS-DA, partial least squares discriminant analysis; VIP, variable importance for the projection; FC, fold change

1. INTRODUCTION

Chinese herbal injection (CHI), which is produced under modern scientific theory to extract and purify effective components from herbs, plays a significant role in the clinical treatment of severe diseases in China. Benefiting from the high bioavailability and rapid curative effect, more than 140 types of CHIs are currently used for more than 400 million patients annually in Chinese hospitals (Li et al., 2018). However, compared with other Chinese medicine formulations, the incidence of adverse drug reaction is obviously high due to the use of CHI (Wang et al., 2009). Allergic reaction (AR), which is one of the most frequent, acute and potentially fatal adverse drug reactions, may lead to great physical and mental damage as well as lay a huge financial burden on patients. So far, the AR of CHI has become one of the most noticeable focuses of public health in China.

The quality of CHI is considered as a potential risk factor for inducing AR. However, the research of quality factor directly related to AR is extremely rare in the literatures. In China, the reevaluation of CHI has been proposed for more than ten years, but little substantial progress has been made due to the lack of methodology (Ren et al., 2011). The current methods primarily focus on bioactive components (Zhang et al., 2018 Wu et al., 2018 Jiang et al., 2018 Zhuo et al., 2017 Jiang et al., 2017 Li et al., 2017 Shi et al., 2017 Jiang et al., 2018). This has been a big problem in improving the clinical safety since other substances are attributed to the cause of AR. So far, the sensitization of some macromolecules like specific proteins is clear-cut. Nevertheless, it still has been remained considerable controversial whether low-molecular-weight components have potential sensitization.

According to statistics, more than 90% of CHI-induced AR cases were related to CHI containing organic acids, and most clinicians considered organic acids as the causes of allergic diseases (Lin et al., 2012). In addition, an intravenous mouse model showed that organic acids were potential sensitizers which triggered immune-mediated drug hypersensitivity by significant enhancing the secretion of IgE (Lin et al., 2012). On the other hand, some views suggested that organic acids mitigated hypersensitivity via oral or subcutaneous exposure (Anderson et al., 2020 Pragasa et al., 2013). With respect to these

studies, the evaluation result of sensitization of organic acids seems to be inconsistent. Therefore, more studies are needed to explore the influence of low-molecular-weight components especially organic acids on AR of CHI.

On account of species difference, the clinical manifestation of human beings is supposed to be an ideal source of AR information. Traditional clinical trial and real-world evidence are both powerful means of obtaining AR information. Traditional clinical trial can take measures designed to control variability and to ensure the quality of the data. However, it is often conducted with specific populations and in specialized environments that differ from the realities of clinical settings. Real-world evidence, which can be obtained from abundant sources efficiently saving time and money, has the potential to generalize findings that reflect the realities of clinical care yielding answers relevant to broader populations of patients than would be possible in a specialized research environment (Sherman et al., 2016). As the expense of conducting traditional clinical trial growing, many clinical researchers have become increasingly interested into real-world evidence (Wang et al., 2018 Losada et al., 2018). In our previous work, the method for obtaining adverse drug reaction information of CHI has been established by detailed reviewing and comprehensive analyzing clinical real-world evidence (Jiang et al., 2019). Nevertheless, no available study for clarifying the relationship between AR information based on real-world evidence and low-molecular-weight components of CHI has been reported to the best of our knowledge.

Shenmai injection, made from *Radix ginseng* and *Radix Ophiopogonis*, is one of the most commonly used CHIs in China. The quality of Shenmai injection has been extensively investigated focusing on bioactive components like ginsenosides and ophioponins using HPLC, HPLC-MS and HPLC-MS/MS (Fang et al., 2006 Zhang et al., 2003 Jiang et al., 2013). However, the influences of low-molecular-weight components on AR are still unclear. Ultra performance liquid chromatography-quadrupole time-of-flight mass spectroscopy (UPLC-Q-TOF-MS), which allows efficient separation, high sensitive and simultaneous analysis of thousands of components, is considered as one of the most effective analytical technologies for low-molecular-weight components mapping (Li et al., 2018). It has been successfully applied to discriminating species, cultivation ages and genera of Chinese medicines (Kim et al., 2011 Shin et al., 2016 Park et al., 2014 Guo et al., 2019 Zhao et al.,

2018). It seems worth a try to reveal the differences in low-molecular-weight component profile of Shenmai injection with different AR risks by combining real-world evidence and UPLC-Q-TOF-MS.

In this study, the relationship between AR and low-molecular-weight component profile of Shenmai injection was explored by an interdisciplinary technology integrating real-world evidence and UPLC-Q-TOF-MS. The approach for analyzing real-world evidence of hospitalized patients was developed to obtain AR information. The UPLC-Q-TOF-MS was exploited to systematically investigate low-molecular-weight component profile with 50-1500 *m/z* mass range. The partial least squares discriminant analysis (PLS-DA) model was then established to map the influence of low-molecular-weight components on AR. Finally, influential components were tentatively identified and the potential sensitization of organic acids was revealed. By using the method, the Shenmai injection with high AR risk could be precisely recognized before clinical use.

2. EXPERIMENTAL

2.1. Sample materials and chemicals

This real-world study of Shenmai injection (50 mL, Chiatai Qingchunbao Pharmaceutical Co., Ltd., Hangzhou, China) was conducted among adult hospitalized patients from January 2015 to June 2019 at two hospitals in China. The cases from Tongde hospital of Zhejiang province, which is an integrated Chinese and western medicine hospital with a total capacity of 1800 licensed beds, were utilized as calibration cohort. While the cases from Zhejiang province hospital of traditional Chinese medicine, which is a 2000-beds traditional Chinese medicine hospital, were utilized as external validation cohort.

In two hospitals, the clinical manifestations of hospitalized patients during and after using Shenmai injection were monitored. The criteria for evaluation AR were basically consistent with Technology Guidelines on Studies for the Immune Toxicities in Chinese Medicine and Natural Medicine. If suspected AR was observed, the real-world evidence of each suspected case was detailed reviewed and comprehensive analyzed through several steps. Firstly, the case with improper or specific medication was filtered out after analyzing

medication information (daily dose, frequency, solvent type, et al.). Secondly, the relevance between AR and the use of Shenmai injection was assessed by six researchers independently through analyzing AR information (occurrence time, clinical symptoms, whether the symptom was relieved or vanished after decreasing dosage or drug withdrawal, et al.). According to the relevance evaluation method of the national adverse drug reaction monitoring center, the suspected AR cases were classified into six grades, including definite, probable, possible, unlikely, pending and unassessable. Only all the six researchers met to reach a consensus decision that the AR had definite or probable relation with the use of Shenmai injection, the case was considered as confirmed AR case. Thirdly, the batch number of Shenmai injection in each confirmed AR case was extracted from hospital drugstore information system and adverse drug reaction monitoring system separately, and then checked at pharmacy intravenous admixture services or inpatient ward. Only when the batch number had been reconfirmed, the batch was incorporated into AR group.

Meanwhile, the clinical manifestations of other patients were monitored. If no AR had been observed until the end of treatment, the batches were involved in control group. In addition, the clinical usage of each batch in control group was extracted from the hospital drugstore information system. The batch was excluded if it had not been used by sufficient patients according to the rate of suspected AR.

HPLC grade acetonitrile was purchased from Merck (Darmstadt, Germany). Formic acid was purchased from Sigma-Aldrich (Darmstadt, Germany). Distilled water was purchased from Watson's Food & Beverage Co., Ltd. (Guangzhou, China).

2.2. UPLC-Q-TOF-MS analysis

The ExionLC UPLC system (AB SCIEX, USA) equipped with a binary pump, a vacuumdegasser and an autosampler was controlled by OS software (AB SCIEX, USA). A Waters Acquity UPLC BEH C18 column (2.1×150 mm, 1.7 μm) was used to perform the chromatographic separation. The autosampler and column temperature was maintained at 4 °C and 40 °C, respectively. The mobile phase consisted of solvent A (0.1 % formic acid-acetonitrile) and solvent B (0.1 % formic acid-water) was delivered at a flow rate of 0.3

mL·min⁻¹. The gradient program used was as follows: 0-2 min, 5-30% A; 2-10 min, 30-42% A; 10-14 min, 42-99% A; 14-16 min, 99% A.

The X-500R Q-TOF-MS (AB SCIEX, USA) was operated in full scan TOF-MS and MS/MS modes with information dependent acquisition (IDA) in a single run analysis. Electrospray ionization (ESI) negative ion mode was used for each sample. TOF-MS was operated in the 50-1500 *m/z* mass range of small molecule mode with dynamic background subtraction setting with the following parameters: ion source gas1 and gas2 of 55 psi, source temperature of 600 °C, curtain gas of 35 psi, CAD gas of 7 psi, spray voltage of -4500 V, declustering potential of -60 V, collision energy of -10 V, accumulation time of 150 ms, maximum candidate ions of 15 and intensity threshold exceeds of 100 cps. MS/MS was programmed in a mass range of 50-1000 *m/z* with accumulation time of 35 ms, declustering potential of -60 V and collision energy of -35±15 V.

80 µL of Shenmai injection was diluted to 1 mL with distilled water and centrifuged at 14000 rpm for 15 min. For low-molecular-weight components profiling, 2 µL of each supernatant was injected for UPLC-Q-TOF-MS analysis.

2.3. Multivariate analysis

To reduce heteroscedasticity, the low-molecular-weight component profile was preprocessed for dealing with zero values. Peaks having larger than 50 percent of the number of zero values across all samples were eliminated. The remaining zero values were replaced with half of the minimum positive values in the original data assuming that the zero values were caused by low abundance components.

Three methods, including PLS-DA, fold change (FC) analysis and t-test, were used to select influential variables. The variables, which met the criterion of each method, were selected to establish PLS-DA models separately. The data set was scaled by unit variance for unifying the influence of each variable to make individual variables more comparable. The influential variables were selected by comparing the predictive ability using Q2 with 7-fold cross-validation.

The optimized PLS-DA model was established based on influential variables. On the basis of the model, the software was devised to provide a convenient tool for application. The samples with different AR risks were then discriminated using the software. The Shenmai injections with predicted values higher than the threshold were recognized as high risk samples. The recognition accuracy for both AR and control group in calibration set were conducted for evaluating the effectiveness of the method. The predictive capacity of the method for new samples was further evaluated through its application to external validation set. The PLS-DA algorithm was performed using SIMCA-P+ (Umetrics, Sweden, version 13.0). The FC analysis and t-test were performed using the MetaboAnalyst 4.0 online (www.metaboanalyst.ca/faces/home.xhtml). Visual Basic (Microsoft, USA, version 6.0) was used as software development environment.

2.4. Ethics, consent and permissions.

The study was approved by the ethics committee of Tongde hospital of Zhejiang province and Zhejiang province hospital of traditional Chinese medicine. This study was performed in accordance with the reporting guidelines and Helsinki declaration. Informed consent was waived due to retrospective design of this study after being approved by the ethics committee.

3. RESULTS

3.1. Description of samples

From January 2015 to June 2019, a total of 8,581 patients used Shenmai injection in the first hospital, among which 64 patients had suspected AR. 14 cases with improper usage and 1 case with inconsistent decisions about the relevance were filtered out. The 49 consistent cases were classified as 3 “definite”, 33 “probable” and 13 “possible”. The 13 “possible” cases were excluded since the probability of AR caused by combined drugs or other treatment could not be ruled out. In 36 confirmed AR cases, most cases were severe according to Technology Guidelines on Studies for the Immune Toxicities in Chinese Medicine and

Natural Medicine. The seriousness and main manifestations of AR in calibration cohort are outlined in Table 1.

After extracting and checking the batch numbers, 26 and 24 samples were categorized as AR and control group. The flowchart of sample selection in calibration set is shown in Figure 1. Using the same method, 10 and 9 samples from the other hospital were discontinuously collected and categorized as AR and control group in external validation set. Finally, 50 samples from the first hospital were utilized as calibration set for selecting influential variables and establishing model, while 19 samples from the second hospital were utilized as the external validation set for validating the predictive capacity of the method.

3.2. Overview of low-molecular-weight component profile

A total of 3,725 MS1 peaks were detected from Shenmai injection by UPLC-Q-TOF-MS. A representative total ion chromatogram and 2D ion intensity map of Shenmai injection is shown in Figure 2A-B. No variable was removed for 50 percent threshold. To obtain an overview of low-molecular-weight component profile, the preliminary PLS-DA model was established based on the data set using all variables. The R^2Y and Q^2 was 0.969 and 0.472, respectively. The scores plot of preliminary PLS-DA model is shown in Figure 2C. The AR and control group tended to cluster in different regions of the scores plot, indicating that there were different characteristics of low-molecular-weight component profile between two groups.

3.3. Selection of influential variables

Although 3,725 MS1 peaks were detected, only some were influential in discrimination; others might even interrupt the identification and should be eliminated (Kim et al., 2011). To obtain a desirable result in the discrimination of new samples, three methods including PLS-DA, FC analysis and t-test were used to identifying the relevant variables. The PLS-DA was applied to evaluating each contribution of the variable to the classification using Variable Importance in Projection (VIP) values. The FC analysis and t-test were conducted to identify the variables having significant differences in two groups using FC and P values, respectively.

The 3D plot, a combination of VIP, FC and P values, is shown in Figure 2D. The number of variables and Q² based on different thresholds of each method are outlined in Figure 3. From 3,725 variables, the PLS-DA, FC analysis and t-test chose 169, 598 and 103 variables, respectively. The optimized Q² of each method was 0.855, 0.771 and 0.832. As it was expected, there were higher Q² values compared with the preliminary model. Finally, 169 influential variables meeting the criterion of VIP value higher than 1.8 were chosen for further classification.

3.4. Establishment and validation of method

The optimized PLS-DA model based on 169 influential variables was established for discrimination analysis. The R²_Y and Q² values were 0.943 and 0.855, respectively. The R²_Y close to 1 and Q² larger than 0.5 indicated good fit of calibration and high predictive ability of cross validation. To further assess the risk that the model was spurious, the R²_Y and Q² of the original model and several models based on data where the order of the group had been randomly permuted were compared. The permutation plot is shown in Figure 4A. All R²_Y and Q² values to the left were lower than values of the original model and the regression line of the Q² points intersected the vertical axis below zero. This result offered further proof that the model fit new samples well.

To provide a convenient tool for AR risk prediction, the software was devised based on the optimized model. The interfaces of the software are shown in Figure 5. With the use of the software, the AR risk of Shenmai injection can be discriminated through three steps. Firstly, double-click to open the software. Secondly, import UPLC-Q-TOF-MS data. Thirdly, click "Predict". The software will substitute the data into the optimized model and recognize the high risk sample automatically within 2 seconds.

The software was applied to predicting AR risk. The predicted value of each sample in calibration set is shown in Figure 4B. The samples with predicted values greater than 0.5 were recognized as high risk ones for inducing AR. All the samples in calibration set were classified correctly.

The effectiveness of the method for new samples was further evaluated through the application of the software to external validation set. The predicted value of each sample in external validation set is shown in Figure 4C. All the samples in external validation set were exactly matched with the true class with 100% accuracy for both AR and control group. The high recognition accuracy indicated the effectiveness of the method in discriminating high risk samples.

3.5. Identification of components

To reveal the relationship between AR and low-molecular-weight components, the structures of components were tentatively identified by comparing their acquired m/z with theoretical data from PubChem online (pubchem.ncbi.nlm.nih.gov) and by comparing their acquired MS/MS spectrum with theoretical data from MassBank online (www.massbank.jp) or traditional Chinese medicine MS/MS library in XCMS online. A total of 27 components were tentatively identified. The qualitative analysis results of 27 components are shown in Table 2. The acquired MS spectrum and acquired/library MS/MS spectrum of 27 components are shown in Figure 6. It was found that the saponins including Ginsenoside Rg1, Ginsenoside Rd, Ginsenoside Rg3, Ginsenoside Rh1, Ophiopogonin D' and Liriope muscari baily saponins C, which are considered to be bioactive in *Radix ginseng* and *Radix Ophiopogonis*, had no significant differences in AR and control group. The levels of citric acid in AR group were significantly higher than those of control group. The representative extracted ion chromatogram, chemical structure and box plot of citric acid are shown in Figure 7. In addition, the levels of protocatechuic acid, p-coumaric acid, quinic acid and malic acid in AR group were higher than those of control group, but the differences were not statistically significant.

4. DISCUSSION

Organic acids, which can be found in many CHIs, have been demonstrated to be therapeutic agents with a variety of biological activities. However, whether organic acids have potential sensitization remains a heatedly debated topic due to the inconsistent results

from the literatures. A meta-analysis indicated that organic acids could induce AR and intravenous administration might be a key factor for sensitization triggering (Lin et al., 2013). The results of this study showed that a positive reaction of potential sensitization could be induced by high levels of citric acid in Shenmai injection. Although the differences of protocatechuic acid, p-coumaric acid, quinic acid and malic acid were not statistically significant, they may also have potential sensitization. To prevent AR, this finding may at least warrant more careful clinical surveillance when using Shenmai injection that contains high levels of organic acids, especially citric acid. To note, the organic acids in other CHIs may also have potential sensitization, which was not covered in this study and clearly needs to be further explored.

In the hapten hypothesis of drug hypersensitivity, the critical step is the formation of adducts (drug-protein conjugates) (Lin et al., 2012). Organic acids are parent components highly reactive toward nucleophiles. Results from cell and animal studies indicated that organic acids induced AR probably through the covalent interactions with macromolecules to form a complete antigen, which stimulated the dendritic cells maturation to launch antigen-presenting (Lin et al., 2012). Deviating from the metabolism of other administration routes like oral or subcutaneous exposure, organic acids enter the bloodstream directly and quickly via intravenous administration. It was speculated that higher levels of organic acids in plasma might be more prone to combining with macromolecules, as a consequence more likely to leading AR.

The PLS-DA, one of the simplest regression algorithms, assumes a certain and simple model that links components with AR. Although many new and complicated algorithms have been developed to capture the relationships more complex than simple linear form, the improvement is limited when tested on new samples (Pencina et al., 2012). It has been demonstrated that the marginal gain from complicated models is typically small compared to the predictive power of the simple models (Pencina et al., 2012). Moreover, the results of some complicated algorithms are hard to explain, which limit their applicability. Therefore, the simple regression model was finally used.

The PLS-DA model creates dummy Y variables for each class and the membership of a class depends upon matching the value of the dummy variable. In this study, the dummy Y

variables were created as 1 for AR group and 0 for control group. Therefore, a predicted value close to 1 indicated membership of AR group and a predicted value close to 0 indicated membership of control group. Finally, 0.5 was used as a threshold in order to classify a sample as belonging to AR group or control group. In addition, different thresholds can be applied according to the practical requirement. For instance, a lower threshold can be set if the criterion for quality controlling is more strict. This means that more samples will be identified as high risk ones and the recognition accuracy in AR group will increased. However, the recognition accuracy in control group will decreased. Likewise, if a relaxer quality controlling criterion is expected, the threshold can be set higher. The recognition accuracy with different thresholds for calibration and external validation set are shown in Figure 4D-E.

With the aim of avoiding poorly conceived study that generated incorrect conclusion, many efforts had been made for analyzing real-world evidence. Firstly, strict exclusion criteria of AR case were applied to increase the power of the model. The cases were excluded if the probability of AR caused by medication factors, combined drugs or other treatment could not be ruled out. Secondly, we tried to reduce bias by strict inclusion criteria of AR case. Only all the six researchers met to reach a consensus decision about the relevance between AR and Shenmai injection, the case was incorporated. Thirdly, strict criteria were set for the selection of samples. Only when the batch number had been repeatedly verified, the batch was included. Finally, 69 samples were involved after analyzing the real-world evidence from thousands of patients for more than four years.

The results of this study must be moderated by the fact that there was limited number of samples. However, the robustness and accuracy of the method had been demonstrated with permuted validation and external validation. In addition, all the cases were non-interventional clinical events, which reflected real-world practice. The completeness, objectivity and accuracy of the AR information can be guaranteed. Despite the inherent limitations, this study has presented the potential sensitization of organic acids in Shenmai injection. For further improving the clinical safety of Shenmai injection, there is a crucial need for larger scale and longer term researches to make a more convincing evidence-bases decision in further works.

5. CONCLUSION

This study presented a method for predicting AR risk of Shenmai injection integrating real-world evidence and UPLC-Q-TOF-MS. The Shenmai injection, which had high AR risk, could be precisely recognized with the application of the devised software. The results of this study suggest that clinical surveillance should be strengthened when using Shenmai injection that contains high levels of organic acids, especially citric acid. We hope that this study will provide a stimulus for future researches to take more attentions to the quality of CHI closely related with clinical safety.

Acknowledgments: This study was supported by the National Natural Science Foundation of China (82104383, 81774270 and 81803808), Zhejiang Provincial Science and Technology Plan of TCM (2018ZY003), Public Service Technology Research Project of Zhejiang Province (2016C33127, LGD19C040005), the New Medical Talent Training Plan of Zhejiang Province in 2017 and the 111 Talent Cultivation Plan of Tongde Hospital of Zhejiang Province (2D01703).

Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

- Anderson K., Ryan N., Siddiqui A., Pero T., Volpedo G., Cooperstone J. L. & Oghumu S. (2020). Black raspberries and protocatechuic acid mitigate DNFB-induced contact hypersensitivity by down-regulating dendritic cell activation and inhibiting mediators of effector responses. *Nutrients*, 12: 1701. <https://doi.org/10.3390/nu12061701>
- Fang X., Wang Y. & Cheng Y. (2006). LC/MS fingerprinting of Shenmai injection: a novel approach to quality control of herbal medicines. *Journal of Pharmaceutical and Biomedical Analysis*, 40: 591-597. <https://doi.org/10.1016/j.jpba.2005.10.036>
- Guo L., Zhang D., Wang L., Xue Z., Guo M., Duan L. & Zheng Y. (2019). Comparison and discrimination of *Artemisia argyi* and *Artemisia lavandulifolia* by gas chromatography-mass spectrometry-based metabolomic approach. *Journal of Aoac International*, 102: 1814-1821. <https://doi.org/10.5740/jaoacint.19-0080>
- Jiang C., Liu Y. & Qu H. (2013). Data fusion strategy based on near infrared spectra and ultraviolet spectra for simultaneous determination of ginsenosides and saccharides in Chinese herbal injection. *Analytical Methods*, 5: 4467-4475. <https://doi.org/10.1039/c3ay26540d>
- Jiang C., Shen J., Shou D., Wang N., Jing J., Zhang G., Gu J., Tian Y., Sun C., He J., Ma J., Wang X. & Li G. (2019). Identification of high-risk patients for ADR induced by traditional Chinese medicine injection: a nested case-control study. *Scientific Reports*, 9: 16721. <https://doi.org/10.1038/s41598-019-53267-2>
- Jiang P., Lu Y. & Chen D. (2017). Qualitative and quantitative analysis of multiple components for quality

- control of Deng-Zhan-Sheng-Mai capsules by ultra high performance liquid chromatography tandem mass spectrometry method coupled with chemometrics. *Journal of Separation Science*, 40: 612-624. <https://doi.org/10.1002/jssc.201600744>
- Jiang Z., Yang J. & Wang Y. (2018). Discrimination and identification of Q-markers based on 'Spider-web' mode for quality control of traditional Chinese medicine. *Phytomedicine*, 44: 98-102. <https://doi.org/10.1016/j.phymed.2017.12.034>
- Jiang Z., Zhao C., Gong X., Sun X., Li H., Zhao Y. & Zhou X. (2018). Quantification and efficient discovery of quality control markers for Emilia Prenanthoidea DC. by fingerprint-efficacy relationship modelling. *Journal of Pharmaceutical and Biomedical Analysis*, 156: 36-44. <https://doi.org/10.1016/j.jpba.2018.04.020>
- Kim N., Kim K., Choi B. Y., Lee D., Shin Y., Bang K., Cha S., Lee J. W., Choi H., Jang D. S. & Lee D. (2011). Metabolomic approach for age discrimination of Panax ginseng using UPLC-Q-ToF MS. *Journal of Agricultural and Food Chemistry*, 59: 10435-10441. <https://doi.org/10.1021/jf201718r>
- Li H., Wang S., Yue Z., Ren X. & Xia J. (2018). Traditional Chinese herbal injection: Current status and future perspectives. *Fitoterapia*, 129: 249-256. <https://doi.org/10.1016/j.fitote.2018.07.009>
- Li M., Du Y., Wang L., Jiang L., Ma X., Zhou P., Li P. & Li H. (2017). Efficient discovery of quality control markers for Gastrodia elata Tuber by fingerprint-efficacy relationship modelling. *Phytochemical Analysis*, 28: 351-359. <https://doi.org/10.1002/pca.2682>
- Li P., Dai W., Lu M., Xie D., Tan J., Yang C., Zhu Y., Lv H., Peng Q., Zhang Y., Guo L., Ni D. & Lin Z. (2018). Metabolomic analysis reveals the composition differences in 13 Chinese tea cultivars of different manufacturing suitabilities. *Journal of the Science of Food and Agriculture*, 98: 1153-1161. <https://doi.org/10.1002/jsfa.8566>
- Lin M., Gong W., Chen Q., Sun L., Wang Y., Fan X. & Shin B. (2013). Evaluation of the potential sensitization of chlorogenic acid: A meta-analysis. *Evidence-Based Complementary and Alternative Medicine*: 208467. <https://doi.org/10.1155/2013/208467>
- Lin M., Gong W., Wang Y., Sun L. & Fan X. (2012). Structure-activity differences of chlorogenic acid and its isomers on sensitization via intravenous exposure. *International Journal of Toxicology*, 31: 602-610. <https://doi.org/10.1177/1091581812461380>
- Lin M., Sun W., Wang Y., Li X., Jin Y., Gong W., Fan X. & Wang Y. (2012). An intravenous exposure mouse model for prediction of potential drug-sensitization using reporter antigens popliteal lymph node assay. *Journal of Applied Toxicology*, 32: 395-401. <https://doi.org/10.1002/jat.1696>
- Losada E., Soldevila B., Ali M. S., Martinez-Laguna D., Nogues X., Puig-Domingo M., Diez-Perez A., Mauricio D. & Prieto-Alhambra D. (2018). Real-world antidiabetic drug use and fracture risk in 12,277 patients with type 2 diabetes mellitus: a nested case-control study. *Osteoporosis International*, 29: 2079-2086. <https://doi.org/10.1007/s00198-018-4581-y>
- Park H., In G., Kim J., Cho B., Han G. & Chang I. (2014). Metabolomic approach for discrimination of processed ginseng genus (Panax ginseng and Panax quinquefolius) using UPLC-QTOF MS. *Journal of Ginseng Research*, 38: 59-65. <https://doi.org/10.1016/j.jgr.2013.11.011>
- Pencina M. J. & Sr D. R. (2012). Thoroughly modern risk prediction? *Science Translational Medicine*, 4: 110f-131f. <https://doi.org/10.1126/scitranslmed.3004127>
- Pragasam S. J., Venkatesan V. & Rasool M. (2013). Immunomodulatory and anti-inflammatory effect of p-coumaric acid, a common dietary polyphenol on experimental inflammation in rats. *Inflammation*, 36: 169-176. <https://doi.org/10.1007/s10753-012-9532-8>
- Ren Y., Zhang P., Yan D., Wang J., Du X & Xiao X. (2011). A strategy for the detection of quality fluctuation of a Chinese herbal injection based on chemical fingerprinting combined with biological fingerprinting.

- Sherman R. E., Anderson S. A., Dal Pan G. J., Gray G. W., Gross T., Hunter N. L., LaVange L., Marinac-Dabic D., Marks P. W., Robb M. A., Shuren J., Temple R., Woodcock J., Yue L. Q. & Califf R. M. (2016). Real-world evidence - What is it and what can it tell us? *New England Journal of Medicine*, 375: 2293-2297. <https://doi.org/10.1056/NEJMs1609216>
- Shi M. & Sun G. (2017). Integrated evaluation of HPLC and UV fingerprints for the quality control of Danshen tablet by systematic quantified fingerprint method combined with antioxidant activity. *Journal of Separation Science*, 40: 1942-1952. <https://doi.org/10.1002/jssc.201601330>
- Shin J., Park H., In G., Seo H. K., Won T. H., Jang K. H., Cho B., Han C. K. & Shin J. (2016). Metabolomic approach for discrimination of four- and six-year-old Red Ginseng (*Panax ginseng*) using UPLC-QToF-MS. *Chemical & Pharmaceutical Bulletin*, 64: 1298-1303. <https://doi.org/10.1248/cpb.c16-00240>
- Wang C., Shi Q., Ding F., Jiang X., Tang W., Yu M. & Zhu J. (2018). Reevaluation of the post-marketing safety of Shuxuening injection based on real-world and evidence-based evaluations. *Drug Design Development and Therapy*, 12: 757-767. <https://doi.org/10.2147/DDDT.S156000>
- Wang H., Ye X., Gao Q., Wu C., Qian Y., Luo B., Sun Y. & He J. (2009). Pharmacovigilance in Traditional Chinese Medicine safety surveillance. *Pharmacoepidemiology and Drug Safety*, 18: 357-361. <https://doi.org/10.1002/pds.1725>
- Wu X., Zhang H., Fan S., Zhang Y., Yang Z., Fan S., Zhuang P. & Zhang Y. (2018). Quality markers based on biological activity: A new strategy for the quality control of traditional Chinese medicine. *Phytomedicine*, 44: 103-108. <https://doi.org/10.1016/j.phymed.2018.01.016>
- Zhang C., Zheng X., Ni H., Li P. & Li H. (2018). Discovery of quality control markers from traditional Chinese medicines by fingerprint-efficacy modeling: Current status and future perspectives. *Journal of Pharmaceutical and Biomedical Analysis*, 159: 296-304. <https://doi.org/10.1016/j.jpba.2018.07.006>
- Zhang H., Wu Y. & Cheng Y. (2003). Analysis of 'shenmai' injection by HPLC/MS/MS. *Journal of Pharmaceutical and Biomedical Analysis*, 31: 175-183. [https://doi.org/10.1016/S0731-7085\(02\)00565-4](https://doi.org/10.1016/S0731-7085(02)00565-4)
- Zhao D., Shi Y., Zhu X., Liu L., Ji P., Long C., Shen Y. & Kennelly E. J. (2018). Identification of potential biomarkers from *Aconitum carmichaelii*, a traditional Chinese medicine, using a metabolomic approach. *Planta Medica*, 84: 434-441. <https://doi.org/10.1055/s-0043-121708>
- Zhuo L., Peng J., Zhao Y., Li D., Xie X., Tong L. & Yu Z. (2017). Screening bioactive quality control markers of QiShenYiQi dripping pills based on the relationship between the ultra-high performance liquid chromatography fingerprint and vascular protective activity. *Journal of Separation Science*, 40: 4076-4084. <https://doi.org/10.1002/jssc.201700514>

Table 1 Seriousness and main manifestations of AR in calibration cohort.

Classification of seriousness	Main manifestations (number of cases*)	Total number of cases*
Mild	Hot flush (5), pruritus (4), erythra (3)	12
Severe	Chest tightness (10), dyspnea (6), palpitation (3), hypotension (2), hyperhidrosis (2), shiver (1)	24
Drug fever	Fever (2)	2
Allergic shock	Pale complexion (0), cyanosis of lips (0), respiratory arrest (0)	0
Other	Sicchasia (3), abdominal pain (1)	4

*The number of cases might be calculated repeatedly since several manifestations might occur simultaneously in one case.

For instance, if hot flush and chest tightness took place simultaneously, the number of cases added 1 to both of them.

Table 2 Qualitative analysis results of 27 components.

N O.	Proposed components	VIP value	FC value	P value	Retention Time/min	Acquired m/z	Molecular formula	Theoretical m/z	Mass Error (ppm)
1	Borneol-2-o-β-d-apiofurano syl(1→6)-β-d-glucopyranoside	2.31	1.07	0.006	5.29	447.2231	C21H36O10	447.2240	-2.01
2	Citric acid	1.83	1.35	0.041	1.47	191.0195	C6H8O7	191.0197	-1.05
3	Protocatechuic acid	1.60	1.10	0.062	2.91	153.0190	C7H6O4	153.0193	-1.96
4	p-coumaric acid	1.47	1.12	0.094	3.92	163.0398	C9H8O3	163.0401	-1.84
5	Liriopeptides B +HCOOH	1.27	1.11	0.15	14.21	767.4211	C39H62O12.H COOH	767.4223	-1.56
6	Quinic acid	1.21	1.07	0.17	1.24	191.0561	C7H12O6	191.0561	0.00
7	Malic acid	1.16	1.04	0.25	1.35	133.0138	C4H6O5	133.0142	-3.01
8	Phenylalanine	1.11	1.06	0.23	2.53	164.0717	C9H11NO2	164.0717	0.00
9	Gypenoside XVII	0.99	1.05	0.51	4.35	945.5432	C48H82O18	945.5428	0.42
10	Notoginsenoside +HCOOH	Ft ₁ 0.98	1.04	0.41	12.31	961.5385	C47H80O17H. COOH	961.5378	0.73
11	Ginsenoside Rg ₁ +HCOOH	0.88	1.01	0.75	4.41	845.4860	C42H72O14.H COOH	845.4904	-5.20
12	Ginsenoside Rd +HCOOH	0.85	1.00	0.98	4.36	991.5447	C48H82O18.H COOH	991.5483	-3.63
13	4-Hydroxybenzoic acid	0.84	1.05	0.36	3.39	137.0244	C7H6O3	137.0244	0.00
14	Neochlorogenic acid	0.77	1.05	0.51	2.86	353.0883	C16H18O9	353.0878	1.42
15	Tryptophan	0.77	1.03	0.84	3.06	203.0825	C11H12N2O2	203.0826	-0.49
16	Gingerglycolipid +HCOOH	B 0.70	0.93	0.47	14.26	723.3786	C33H58O14.H COOH	723.3809	-3.18
17	Linoleic acid	0.65	1.07	0.48	15.34	279.2330	C18H32O2	279.2330	0.00
18	Guanosine	0.53	0.99	0.94	1.92	282.0844	C10H13N5O5	282.0844	0.00
19	Ophiopogonin D' +HCOOH	0.45	1.04	0.64	13.94	899.4638	C44H70O16.H COOH	899.4646	-0.89
20	Adenine	0.44	0.94	0.63	1.88	134.0471	C5H5N5	134.0472	-0.75
21	Chikusetsusaponin IVa	0.42	1.01	0.82	12.98	793.4351	C42H66O14	793.4380	-3.65
22	Ginsenoside Rg ₃ +HCOOH	0.41	1.04	0.72	13.36	829.4945	C42H72O13.H COOH	829.4955	-1.21
23	Ginsenoside Rh ₁ +HCOOH	0.40	0.99	0.90	8.68	683.4371	C36H62O9.HC OOH	683.4376	-0.73
24	Liriope muscari baily saponins C +HCOOH	0.37	1.04	0.68	12.78	915.4592	C44H70O17.H COOH	915.4595	-0.33
25	Caffeic acid	0.37	1.02	0.74	3.50	179.0350	C9H8O4	179.0350	0.00
26	Ginsenoside Rg ₃	0.34	0.99	0.92	13.35	783.4898	C42H72O13	783.4900	-0.26
27	Adenosine	0.14	1.01	0.96	1.88	266.0896	C10H13N5O4	266.0895	0.38

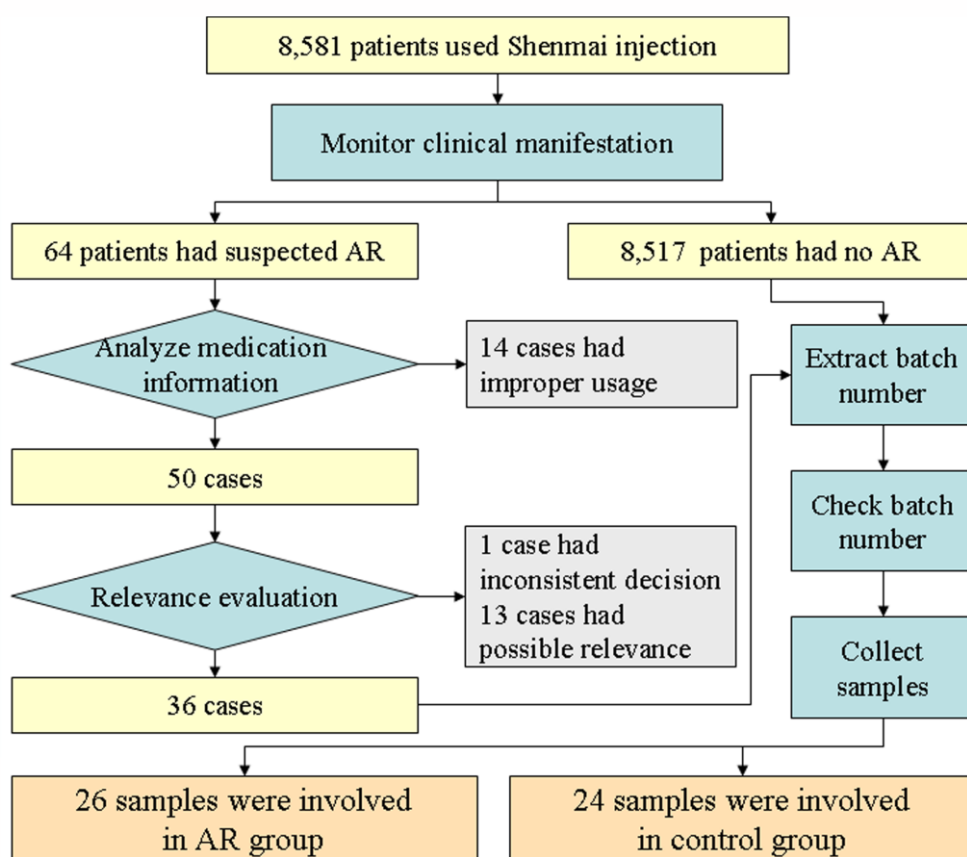


Figure 1. Flowchart of sample selection in calibration set.

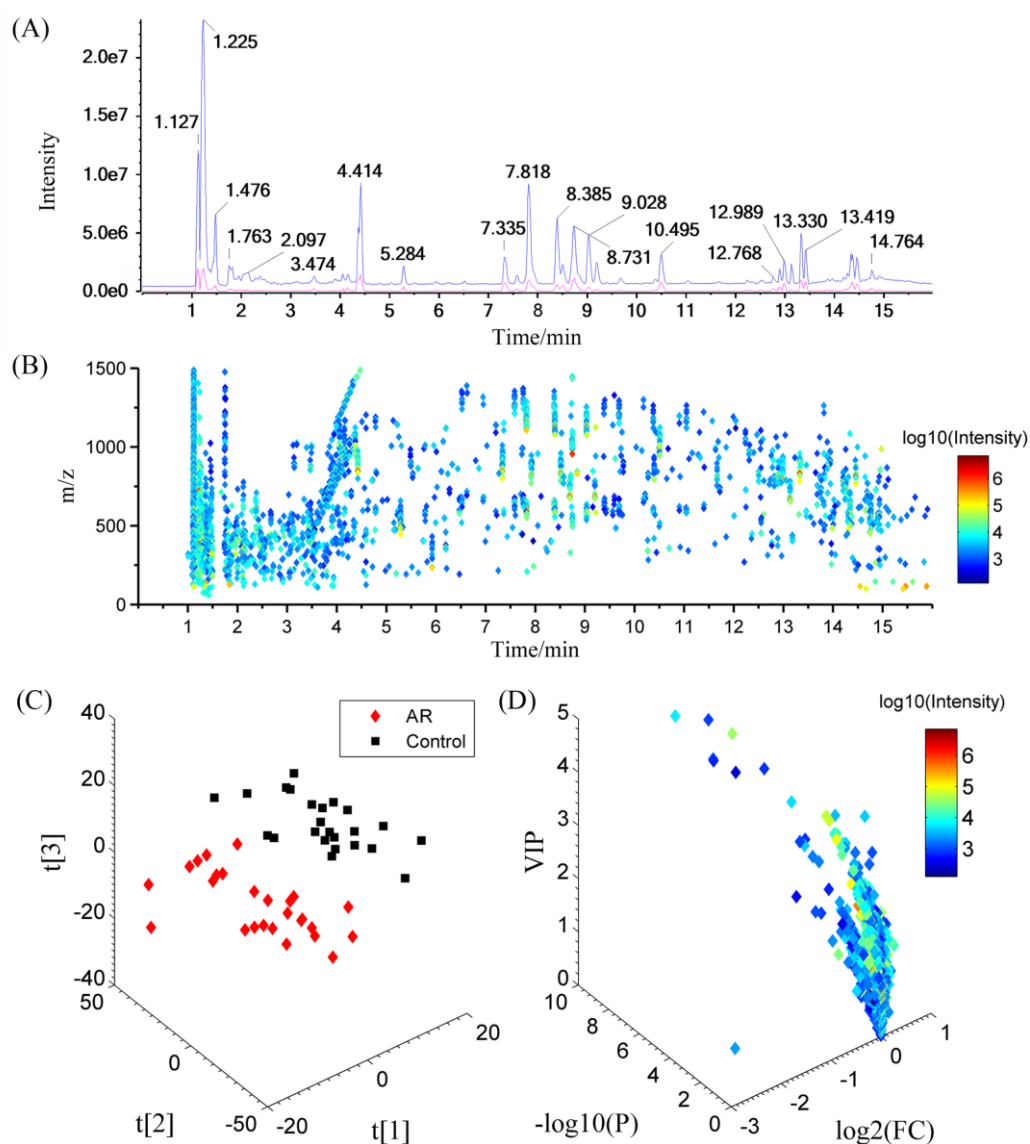


Figure 2. Low-molecular-weight component profile of Shenmai injection. (A) Representative total ion chromatogram. The blue line represents UPLC-Q-TOF-MS total ion chromatogram. The pink line represents UPLC-Q-TOF-MS/MS total ion chromatogram. (B) Representative 2D ion intensity map of 3,725 MS1 peaks. The intensity is log transformed. (C) Scores plot of preliminary PLS-DA model. (D) 3D plot of VIP, FC and P values. The intensity, FC and P values are log transformed.

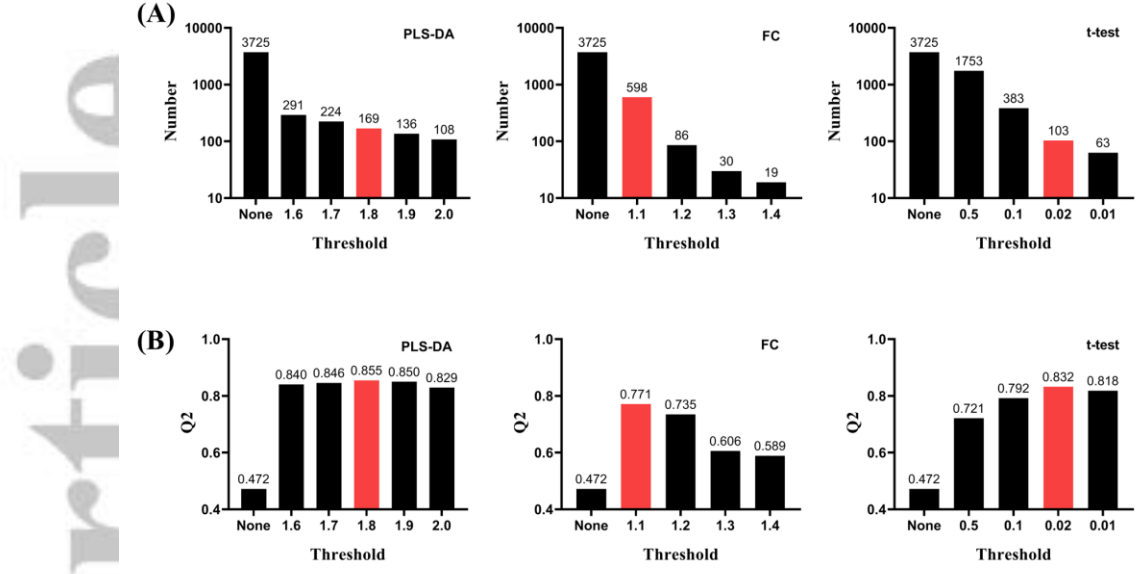


Figure 3. Selection of influential variables based on different thresholds of three methods. (A) Number of variables. (B) Q2. The red bar represents the optimal model of each method.

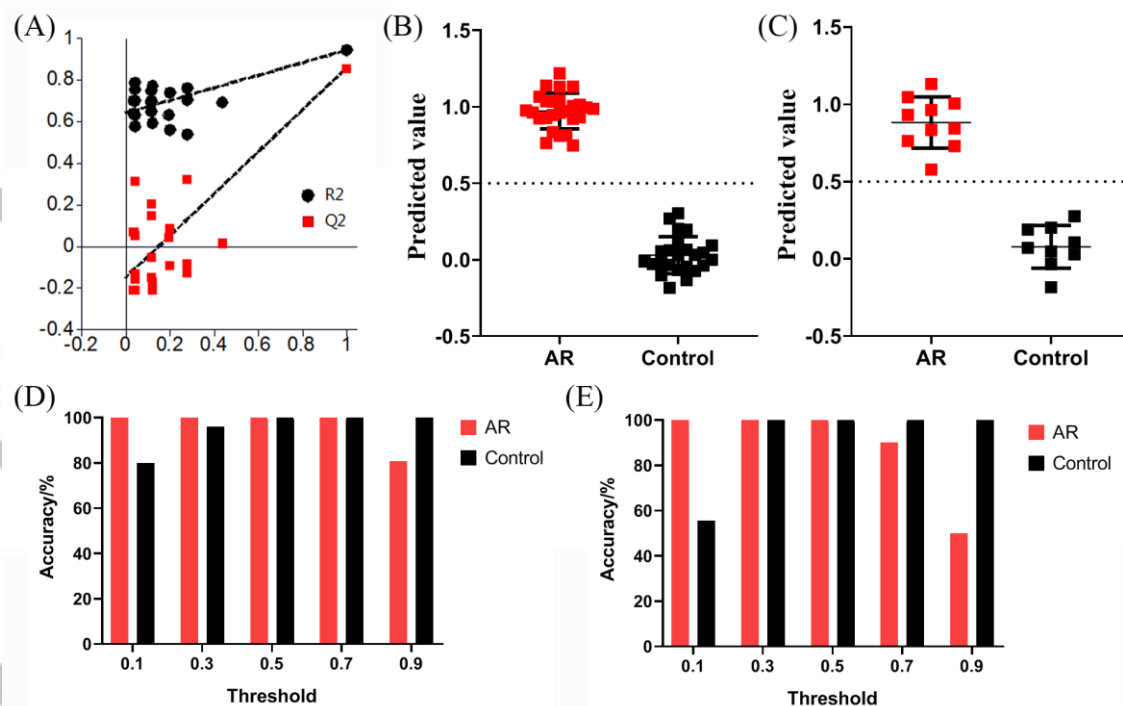


Figure 4. Predicted results of the optimized model. (A) Permutation plot. The number of permutations was 20. (B) Predicted value of each sample in calibration set. (C) Predicted value of each sample in external validation set. (D) Recognition accuracy with different thresholds for calibration set. (E) Recognition accuracy with different thresholds for external validation set.

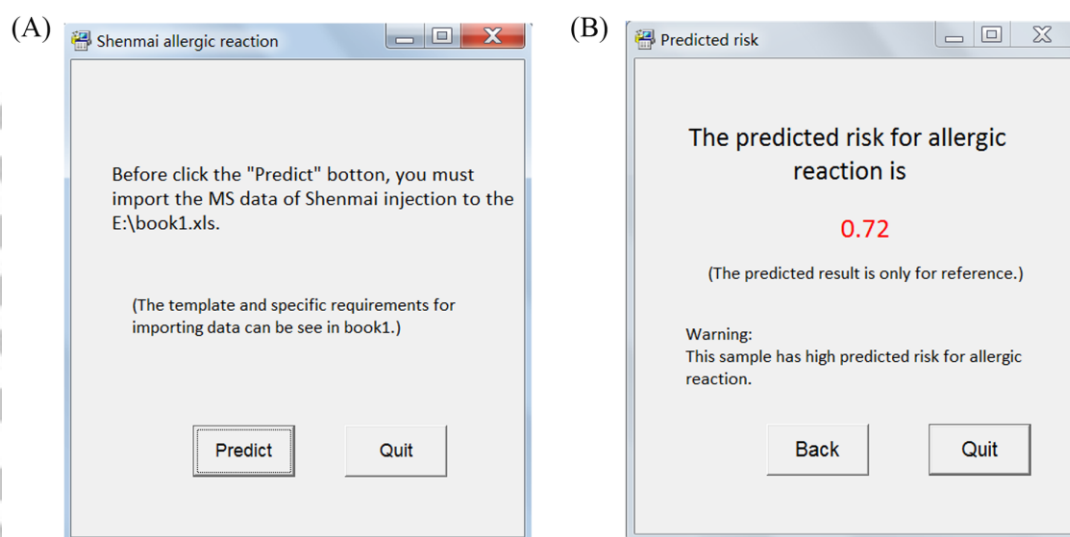
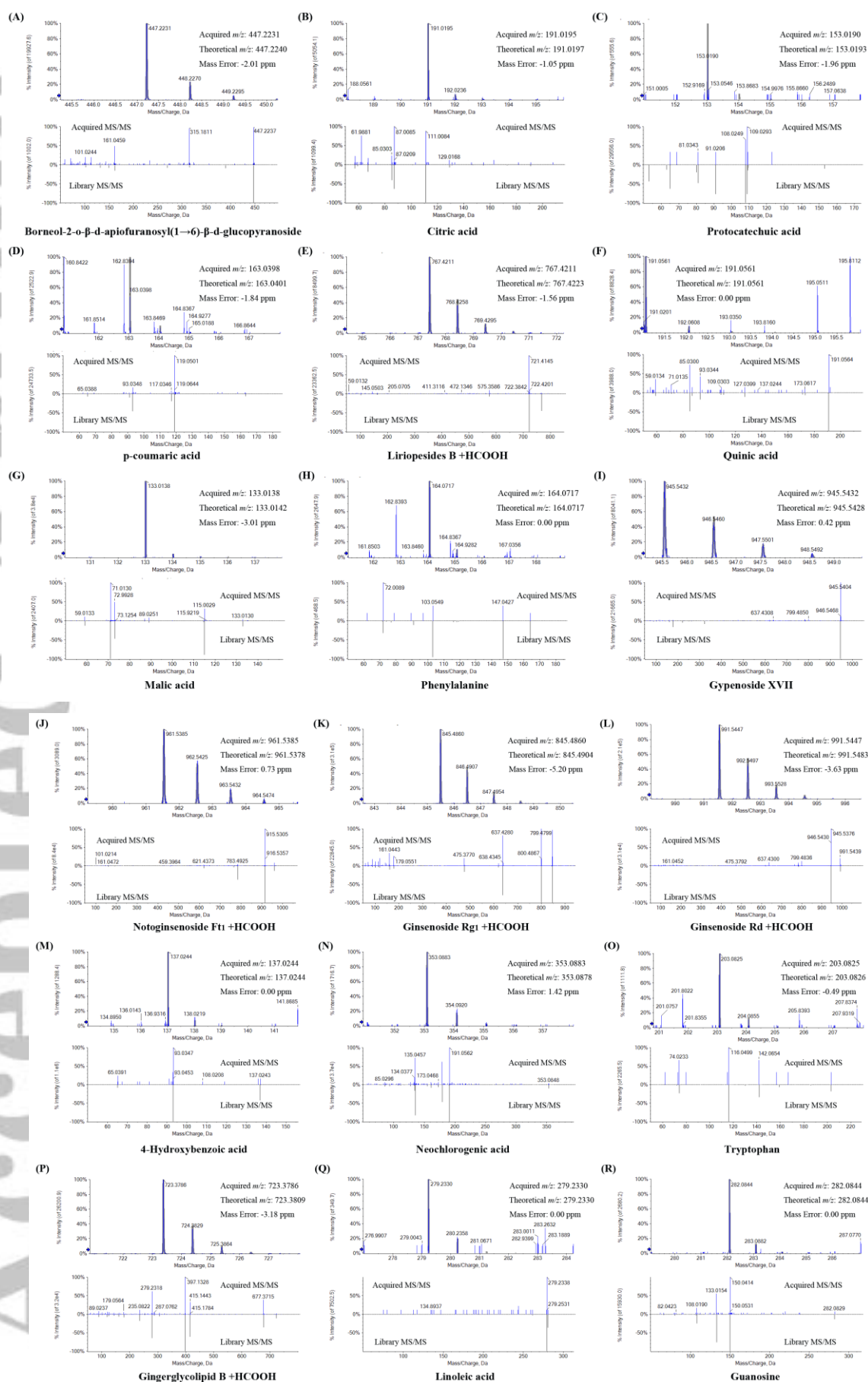


Figure 5. Interfaces of software. (A) Initial interface. (B) Result interface.



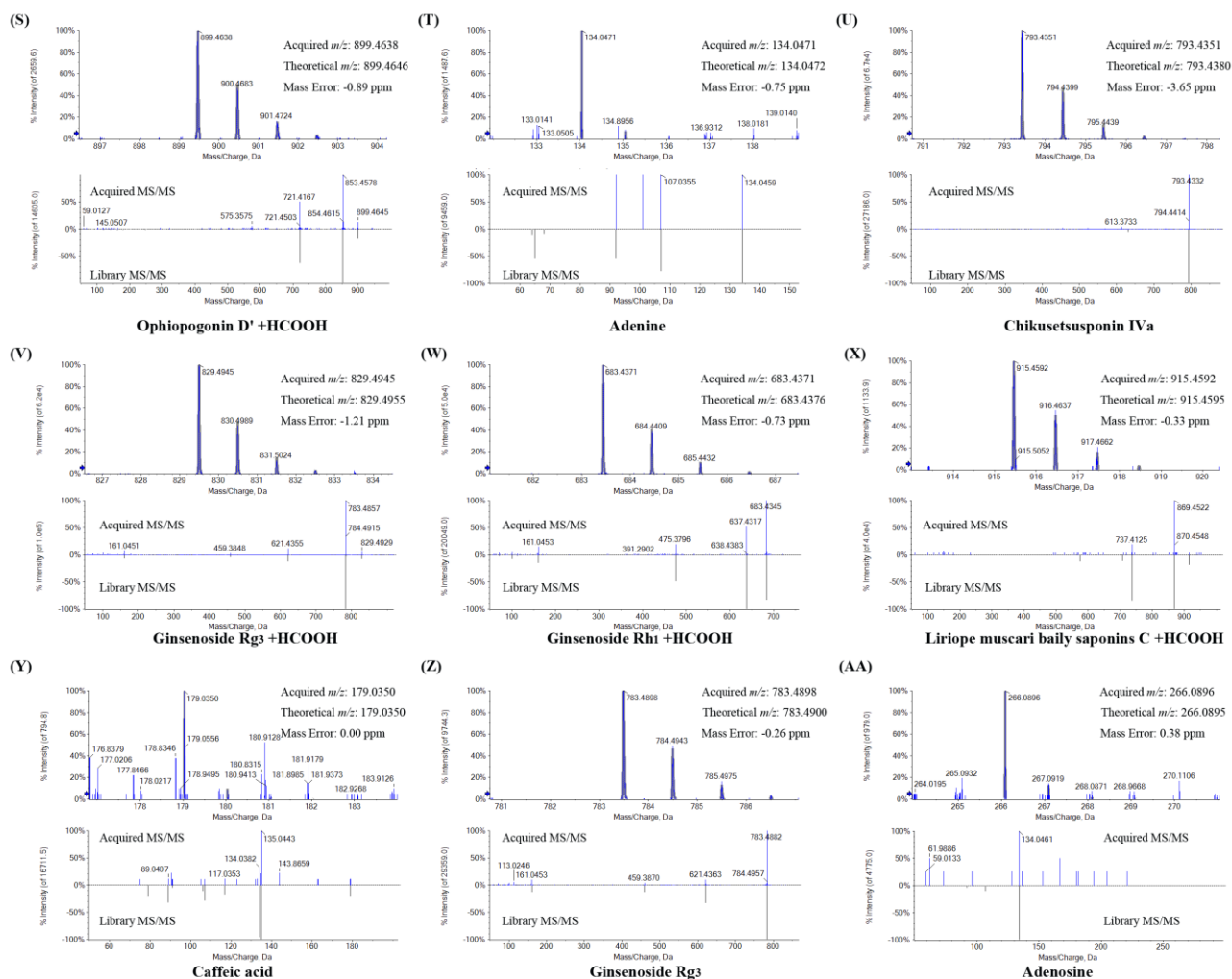


Figure 6. Acquired MS spectrum and acquired/library MS/MS spectrum of 27 components.

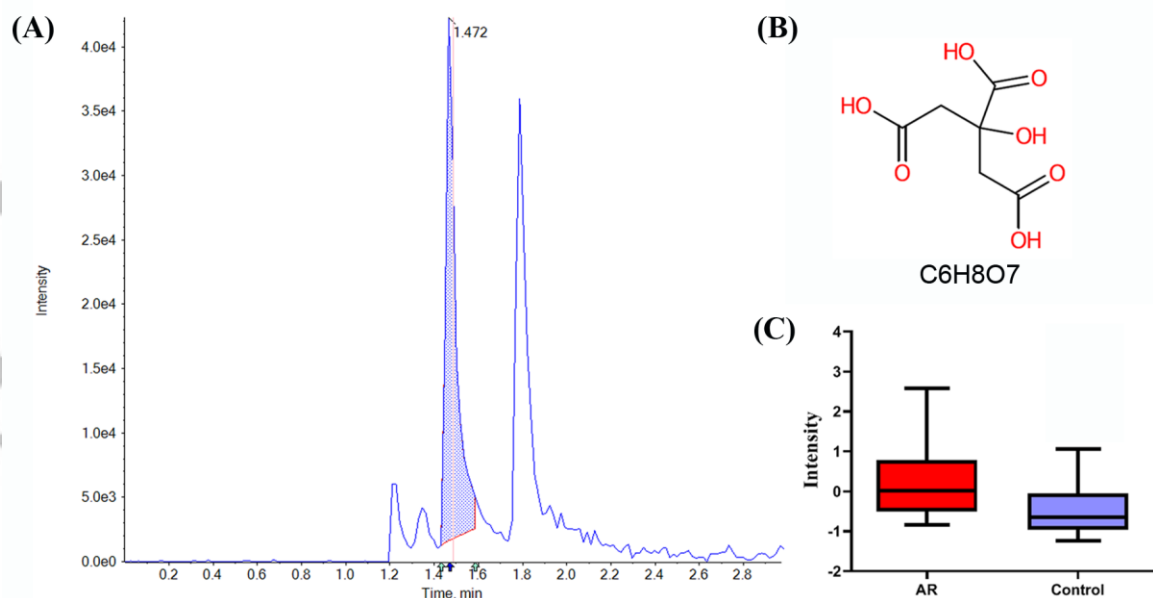


Figure 7. Representative extracted ion chromatogram (A), chemical structure (B) and box plot (C) of citric acid. The intensity in box plot is unit variance scaled.