Supporting Information for

The use of GC-HRMS-based untargeted metabolomics to discover metabolic changes and help in the determination of complex causes of death: A preliminary study

Kai Zhang[1]; Hui Yan[2]; Ruina Liu[1]; Ping Xiang[2]; Ji Zhang[2]; Kaifei Deng[2]; Ping Huang[2] and Zhenyuan Wang[1*]*

[1] Department of Forensic Pathology, College of Forensic Medicine, Xi'an Jiaotong University, Xi'an, 710061, P. R. China;

[2] Shanghai Key Laboratory of Forensic Medicine, Shanghai Forensic Service Platform, Academy of Forensic Science, Shanghai, 200063, P. R. China;

**E-mail address: huangp@ssfjd.cn; wzy218@xjtu.edu.cn*

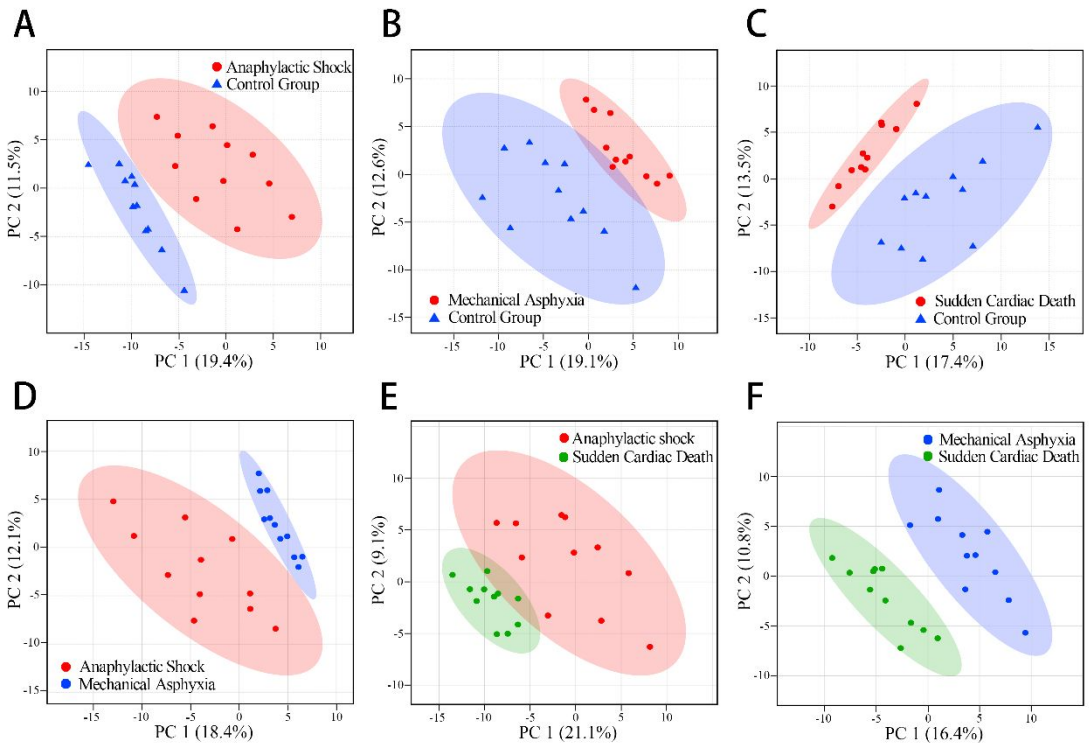

Figure S1. The PCA score plots of the rat plasma GC-HRMS data derived from the three target COD groups and the control group. A: between anaphylactic shock and control group; B: between mechanical asphyxia and control group; C: between sudden cardiac death and control group; D: between anaphylactic shock and mechanical asphyxia group; E: between anaphylactic shock and sudden cardiac death group; F: between mechanical asphyxia and sudden cardiac death group.

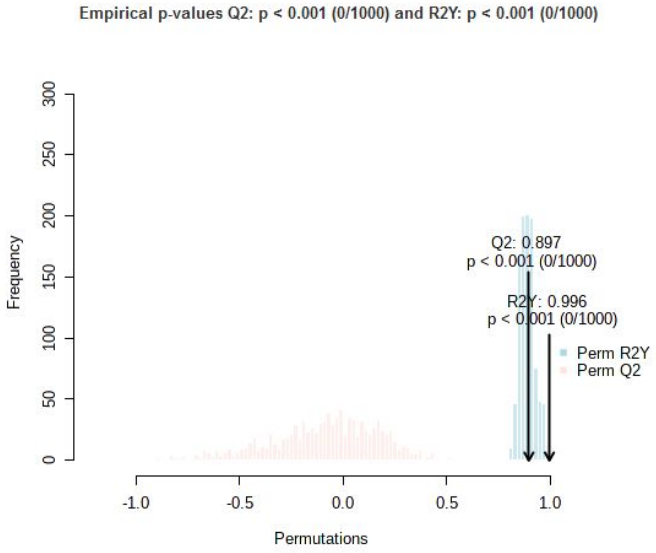

Figure S2. The results from 1000 permutation testing for the PLS-DA model between the anaphylactic shock group and control group.

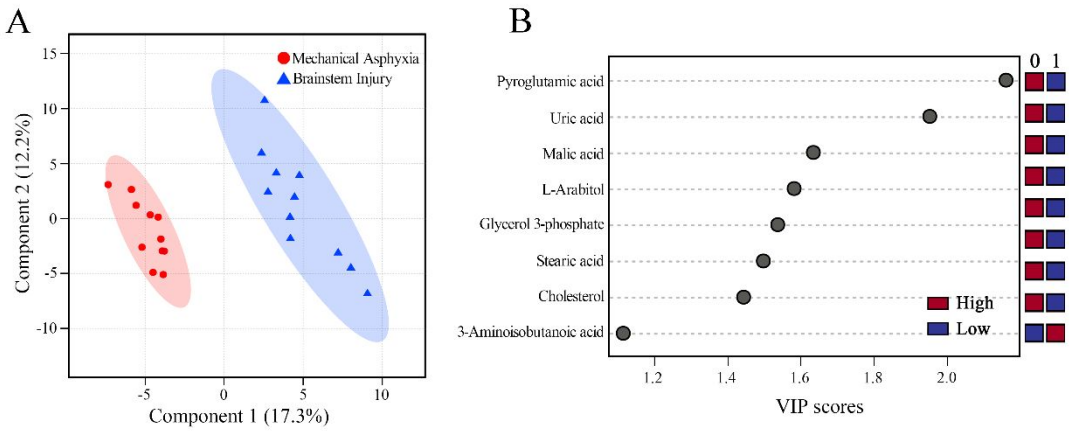

Figure S3. Screening differential metabolites for the mechanical asphyxia group. A: The score plot of the PLS-DA model with R^2 = 0.999 and Q^2 = 0.851. 1000 permutation testing resulted in R^2 = 0.991 and Q^2 = 0.845. B: The VIP score plot of metabolites verified in the PLS-DA model between the mechanical asphyxia group (0) and the control group (1). Red and blue represent the high and low relative contents of the corresponding metabolites, respectively.

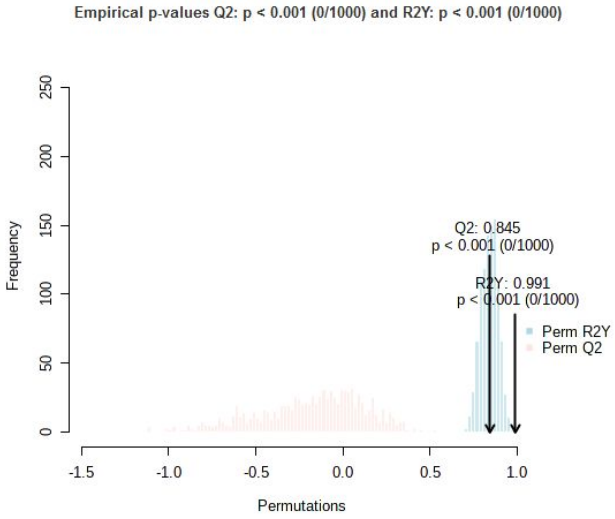

Figure S4. The results from 1000 permutation testing for the PLS-DA model between the asphyxia group and control group.

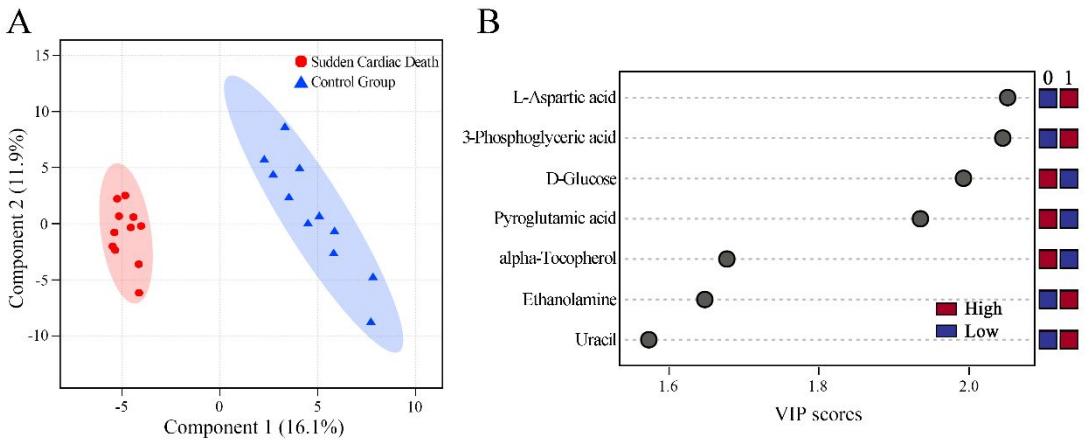

Figure S5. Screening differential metabolites for the sudden cardiac death group. A: The score plot of the PLS-DA model with R^2 = 0.998 and Q^2 = 0.867. 1000 permutation testing resulted in R^2 = 0.978 and Q^2 = 0.857. B: The VIP score plot of metabolites verified in the PLS-DA model between the sudden cardiac death group (0) and the control group (1). Red and blue represent the high and low relative contents of the corresponding metabolites, respectively.

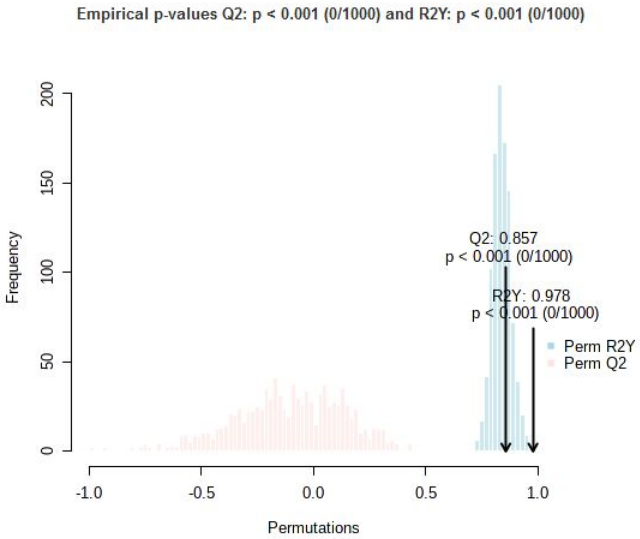

Figure S6. The results from 1000 permutation testing for the PLS-DA model between the sudden cardiac death group and control group.

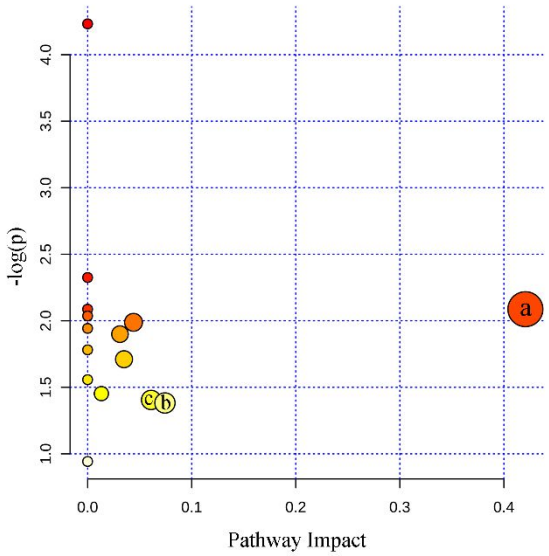

Figure S7. Summary of metabolic pathway of differential metabolites in the anaphylactic asphyxia group. a: Starch and sucrose metabolism; b: Pyrimidine metabolism; c: Arginine and proline metabolism.

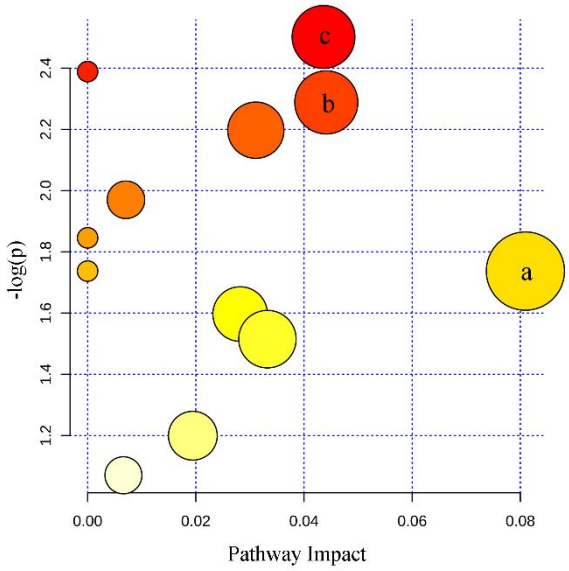

Figure S8. Summary of metabolic pathway of differential metabolites in the mechanical asphyxia group. a: Glycerophospholipid metabolism; b: Citrate cycle (TCA cycle); c: Glycerolipid metabolism.

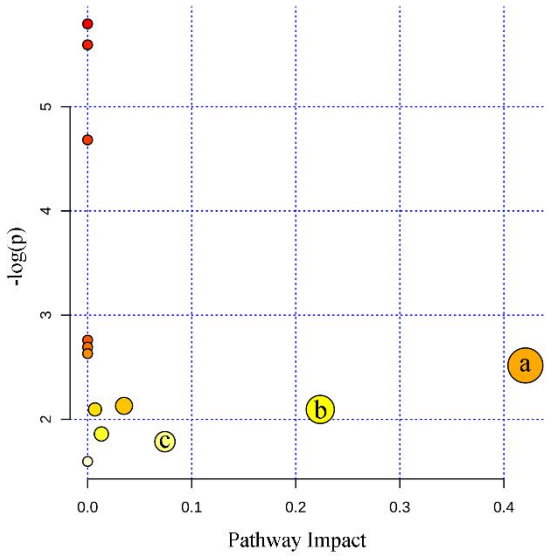

Figure S9. Summary of metabolic pathway of differential metabolites in the sudden cardiac death group. a: Starch and sucrose metabolism; b: Alanine, aspartate and glutamate metabolism; c: Pyrimidine metabolism.

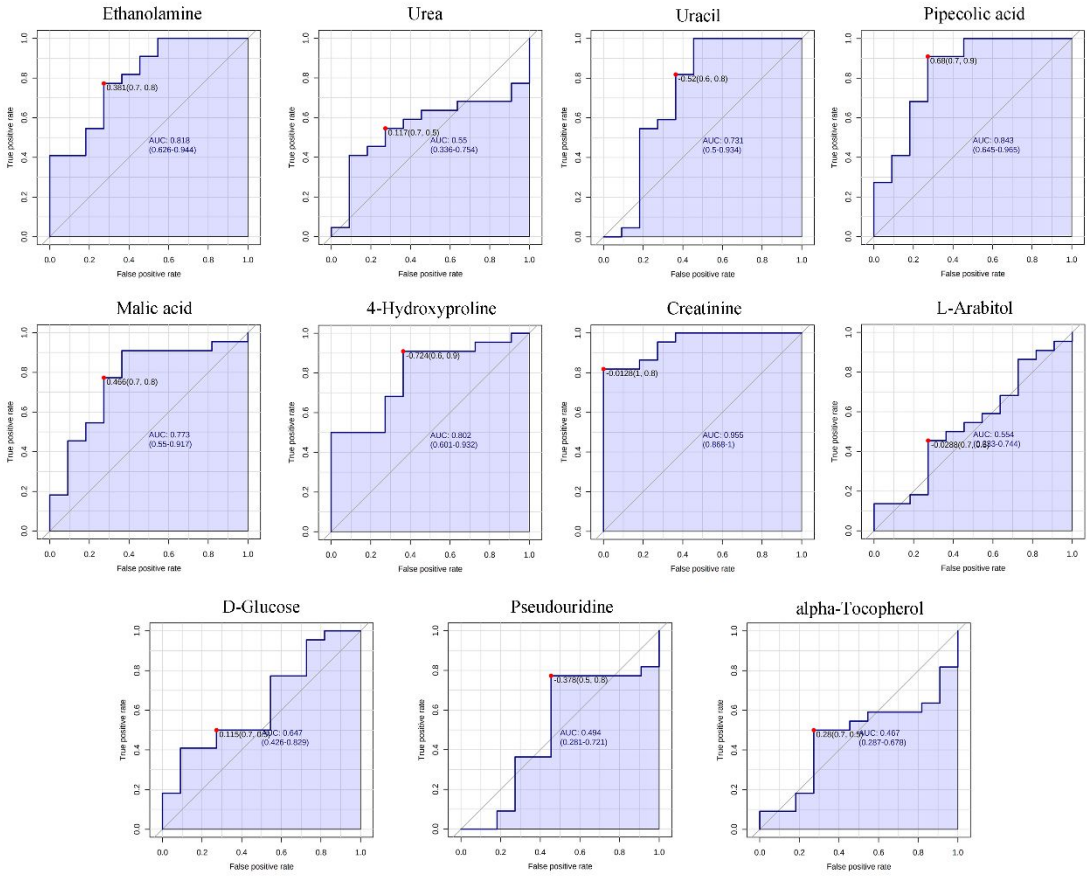

Figure S10. The ROC curves for each verified metabolite in the anaphylactic shock group.

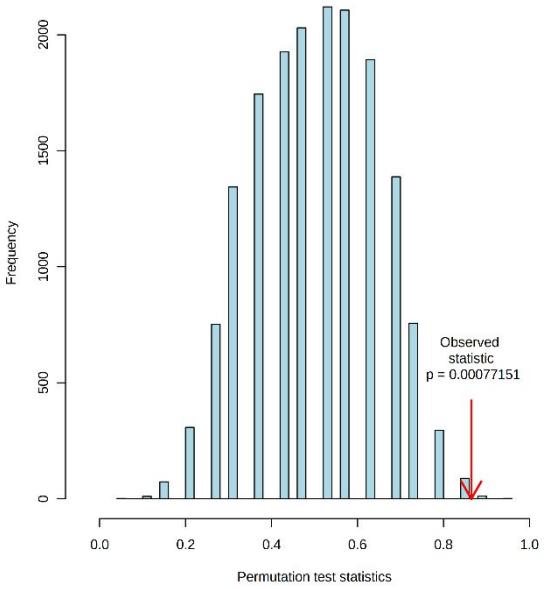

Figure S11. The classification model that used the combined metabolites to discriminate AS and non-AS was evaluated by 1000 permutation testing for the predictive accuracy. The non-permutated model was statistically different from the permutated models ($p < 0.001$), which shows the model against overfitting data.

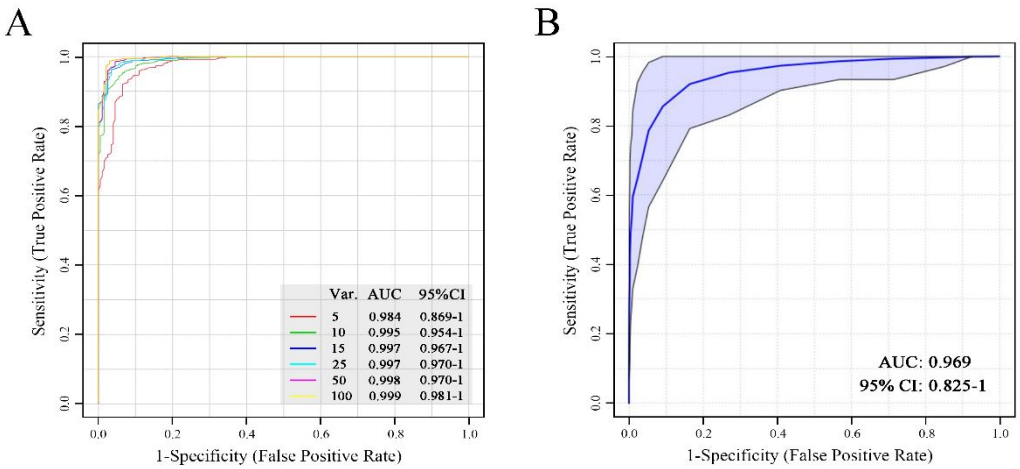

Figure S12. Evaluation of classification ability using all the unselected data (A) and a combination of malic acid, uric acid, 3-aminoisobutanoic acid, and cholesterol (B) via ROC curve analysis for the mechanical asphyxia group. Var.: the number of variables or features used in the model. The six ROC curves were generated using the top 5-100 features ("top" means higher VIP values). AUC: area under the curve.

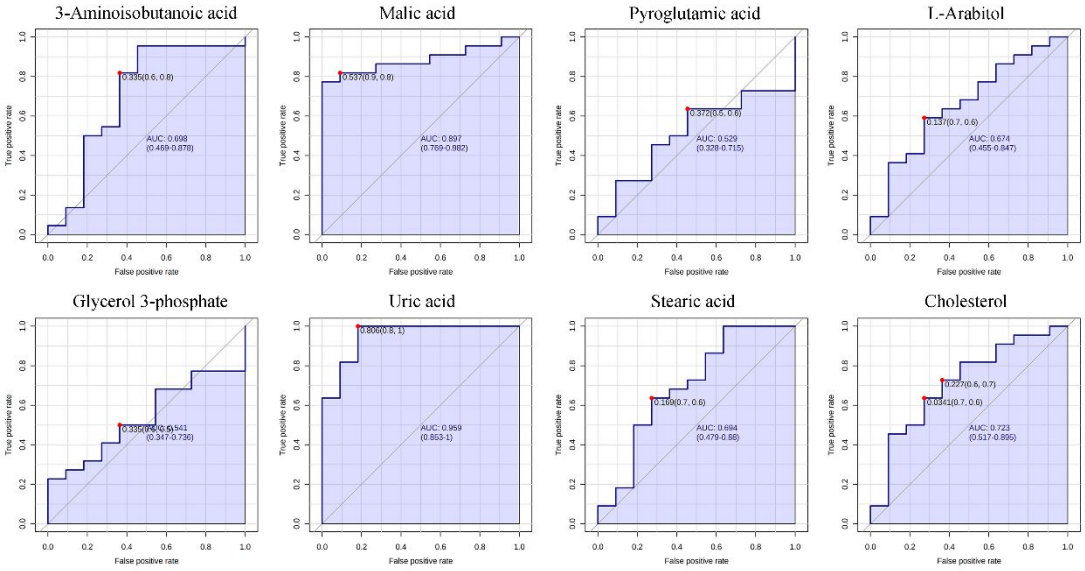

Figure S13. The ROC curves for each verified metabolite in the mechanical asphyxia group.

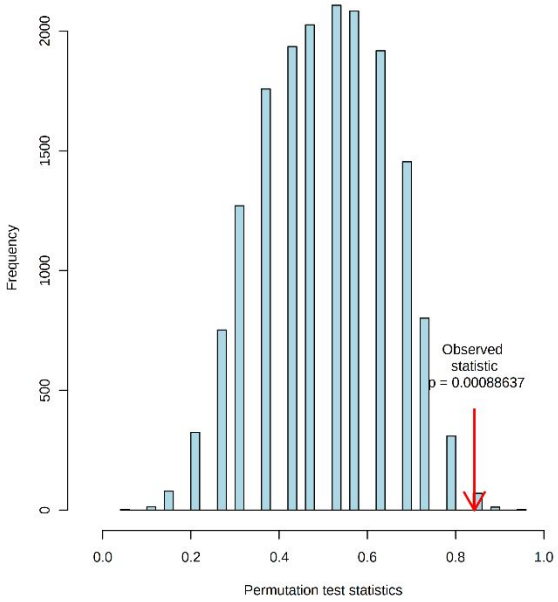

Figure S14. The classification model that used the combined metabolites to discriminate MA and non-MA was evaluated by 1000 permutation testing for the predictive accuracy. The non-permutated model was statistically different from the permutated models ($p < 0.001$), which shows the model against overfitting data.

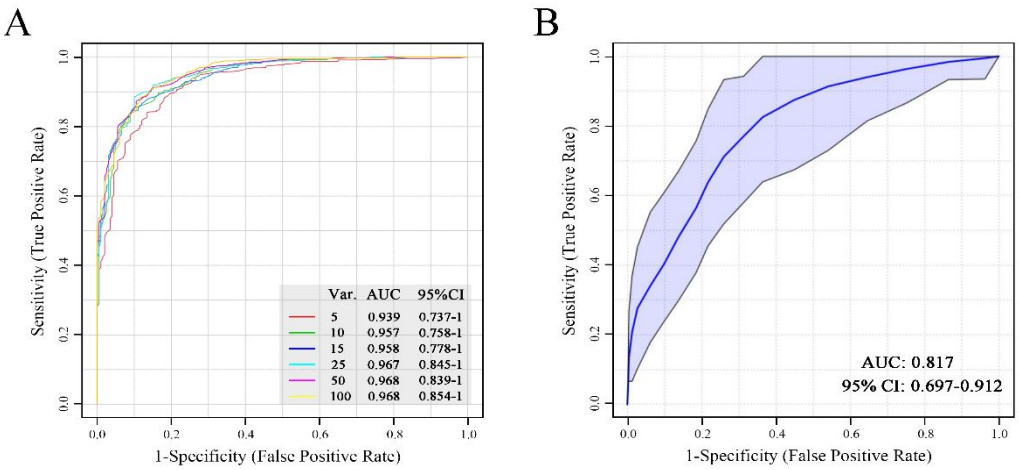

Figure S15. Evaluation of classification ability using all the unselected data (A) and a combination of uracil, pyroglutamic acid, and alpha-tocopherol (B) via ROC curve analysis for the sudden cardiac death group. Var.: the number of variables or features used in the model. The six ROC curves were generated using the top 5-100 features ("top" means higher VIP values). AUC: area under the curve.

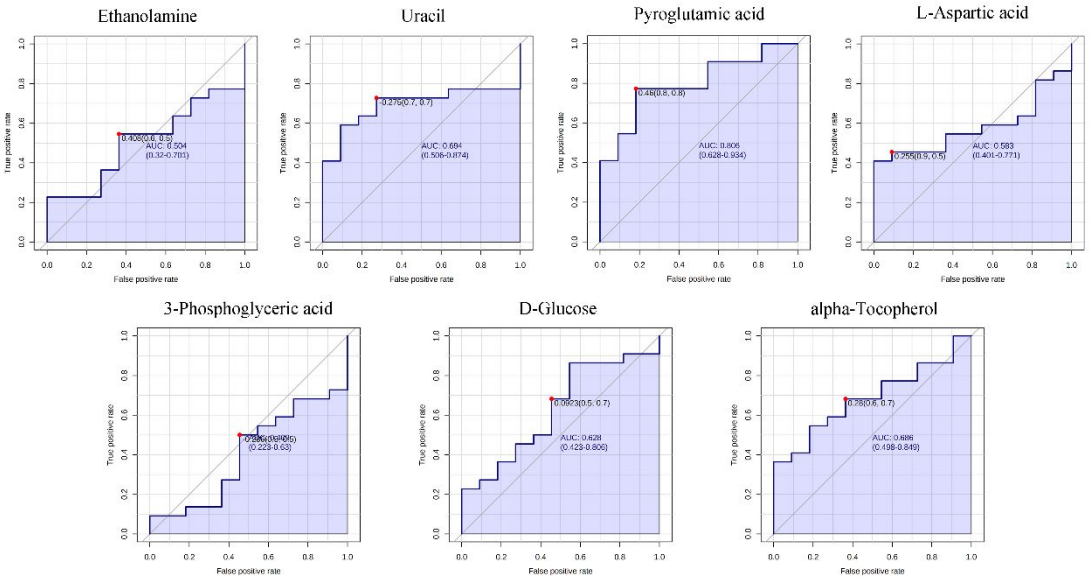

Figure S16. The ROC curves for each verified metabolite in the sudden cardiac death group.

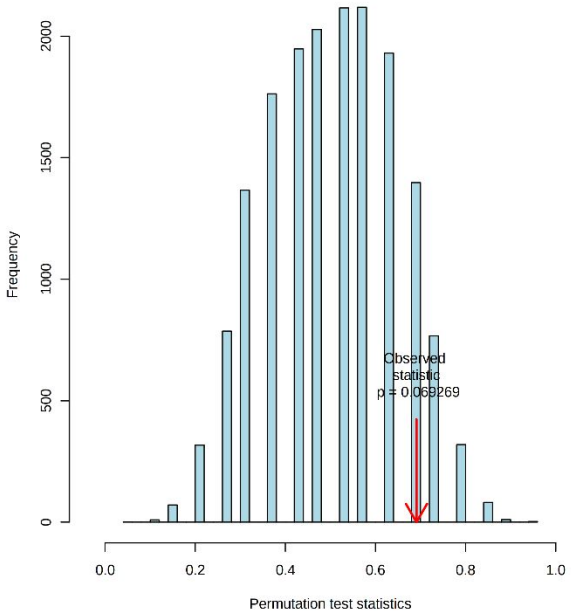

Figure S17. The classification model that used the combined metabolites to discriminate SCD and non-SCD was evaluated by 1000 permutation testing for the predictive accuracy. The p value is greater than 0.05, which shows the risk of overfitting of this classification model.

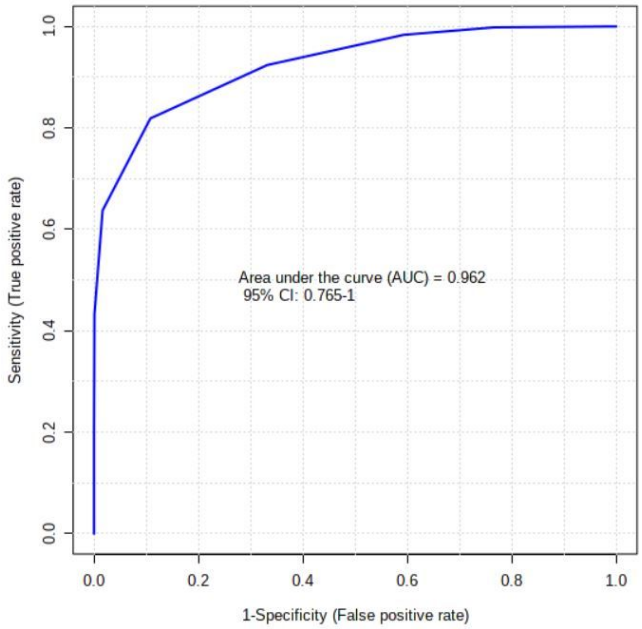

Figure S18. The ROC curve of combination panel (creatinine, pipecolic acid, and ethanolamine) in AS group using data from the test set. CI: confidence interval.

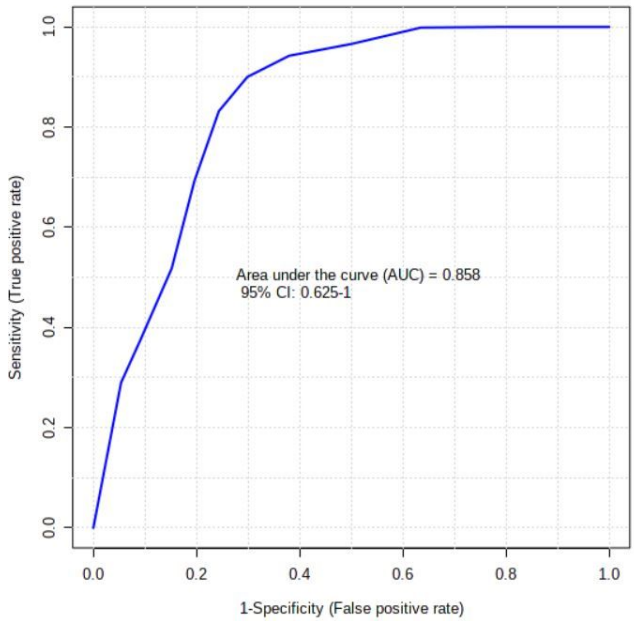

Figure S19. The ROC curve of combination panel (malic acid, uric acid, 3-aminoisobutanoic acid, and cholesterol) in MA group using data from the test set. CI: confidence interval.

Feature	Metabolites	RSD	Feature	Metabolites	RSD
102.07338/4.18	Ethanolamine	0.2877	217.10715/12.89	L-Arabitol	0.1163
147.0655/6.13	Urea	0.2969	299.07083/13.35	Glycrol 3-phosphate	0.1173
102.07338/6.55	3-Aminoisobutanoate	0.2914	299.07074/13.83	3-Phosphoglyceric acid	0.1012
241.08188/8.26	Uracil	0.2419	217.10735/14.52	D-Glucose	0.1895
156.12016/8.65	Pipecolic acid	0.0869	441.16197/16.74	Uric acid	0.0473
147.06557/10.19	Malic acid	0.0998	117.03651/18.08	Stearic acid	0.0840
147.06547/10.61	Pyroglutamic acid	0.2428	217.10716/18.8	Pseudouridine	0.1035
230.13872/10.67	4-Hydroxyproline	0.1079	237.13023/25.14	alpha-Tocopherol	0.2270
232.11801/10.73	L-Aspartic acid	0.0774	129.07288/25.28	Cholesterol	0.0450
115.08111/11.02	Creatinine	0.0851			

Table S1. The relative standard deviation (RSD) of verified differential metabolites. The feature consists of representative m/z and retention time. RSD = standard deviation/average.