

Development of heart failure risk prediction models based on a multi-marker approach using random forest algorithms

Hui Yuan^{1,2}, Xue-Song Fan², Yang Jin², Jian-Xun He², Yuan Gui², Li-Ying Song², Yang Song², Qi Sun², Wei Chen¹

¹Clinical Laboratory, First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shaanxi 710061, China;

²Department of Clinical Laboratory Center, Beijing Anzhen Hospital, Capital Medical University, Beijing Institute of Heart, Lung and Blood Vessel Diseases, Beijing 100029, China.

Abstract

Background: The early identification of heart failure (HF) risk may favorably affect outcomes, and the combination of multiple biomarkers may provide a more comprehensive and valuable means for improving the risk of stratification. This study was conducted to assess the importance of individual cardiac biomarkers creatine kinase MB isoenzyme (CK-MB), B-type natriuretic peptide (BNP), galectin-3 (Gal-3) and soluble suppression of tumorigenicity-2 (sST2) for HF diagnosis, and the predictive performance of the combination of these four biomarkers was analyzed using random forest algorithms.

Methods: A total of 193 participants (80 patients with HF and 113 age- and gender-matched healthy controls) were included from June 2017 to December 2017. The correlation and regression analysis were conducted between cardiac biomarkers and echocardiographic parameters. The accuracy and importance of these predictor variables were assessed using random forest algorithms.

Results: Patients with HF exhibited significantly higher levels of CK-MB, BNP, Gal-3, and sST2. BNP exhibited a good independent predictive capacity for HF (AUC 0.956). However, CK-MB, sST2, and Gal-3 exhibited a modest diagnostic performance for HF, with an AUC of 0.709, 0.711, and 0.777, respectively. BNP was the most important variable, with a remarkably higher mean decrease accuracy and Gini. Furthermore, there was a general increase in predictive performance using the multi-marker model, and the sensitivity, specificity was 91.5% and 96.7%, respectively.

Conclusion: The random forest algorithm provides a robust method to assess the accuracy and importance of predictor variables. The combination of CK-MB, BNP, Gal-3, and sST2 achieves improvement in prediction accuracy for HF.

Keywords: Biomarkers; Diagnostic accuracy; Heart failure; Random forests

Introduction

Heart failure (HF) is a growing public epidemic characterized by ventricular remodeling and variable degrees of myocardial fibrosis, and is associated with significant mortality, morbidity and healthcare expenditures.^[1] Despite the remarkable therapeutic advancement in recent decades, the prognosis of patients with HF remains poor. The early identification of patients with high HF risk may favorably affect outcomes. Recently, there are several biomarkers that have gained mounting interest. B-type natriuretic peptide (BNP) and its amino-terminal fragment (NT-proBNP) are gold standard biomarkers that have been generally used for evaluating patients with HF.^[2] Creatine kinase MB isoenzyme (CK-MB) is a classical indicator for myocardial injury, with excellent sensitivity and specificity.^[3] Suppression of tumorigenicity 2 (ST2) is a receptor expressed in cardiomyocytes and the vascular

endocardium, which has two forms: soluble (sST2) and transmembrane forms. sST2 has been considered to function as a “decoy” receptor for interleukin (IL)-33.^[4] Elevated serum concentrations of sST2 as a biomarker represent disease severity, myocardial stretch and inflammation, as well as unfavorable prognosis.^[5] Galectin-3 (Gal-3) is a soluble β -galactosidase-binding glycoprotein that is secreted by cardiac macrophages. It induces fibroblast proliferation and collagen deposition in the myocardium, and causes ventricular dysfunction as a consequence. An increase in Gal-3 concentration has been found to be associated with increased risk for mortality in patients with chronic HF, regardless of etiology.^[6] The dynamic progression of HF is complex, and is driven by cardiac dysfunction and maladaptive compensatory processes. Given the complexity of this syndrome, it is unlikely that individual biomarkers may be sufficient to assess multi-system disorders, leading to the limited prediction

Access this article online

Quick Response Code:



Website:
www.cmj.org

DOI:
10.1097/CM9.0000000000000149

Correspondence to: Dr. Wei Chen, Clinical Laboratory, First Affiliated Hospital of Xi'an Jiaotong University, No. 277 Yanta West Road, Xi'an, Shaanxi 710061, China
E-Mail: 2398023255@qq.com

Copyright © 2019 The Chinese Medical Association, produced by Wolters Kluwer, Inc. under the CC-BY-NC-ND license. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Chinese Medical Journal 2019;132(7)

Received: 02-11-2018 Edited by: Li-Shao Guo

accuracy of HF. Thus, the combination of multiple biomarkers may provide a more comprehensive and valuable means for improving the risk of stratification.^[7]

The random forest, which was first proposed by Leo Breiman,^[8] is one commonly used ensemble machine learning algorithm for classification, regression and prediction, which operates by building a collection of decision trees at training. The decision tree works by learning simple decision rules extracted from the data features. The output will consider the prediction of each decision tree and make a final decision. Random forests can help identify potential predictors and uncover complex interactions, even in high dimensional settings. This method also simultaneously considers the impact of each individual variable and multivariate interaction with other features. Given its superior performance in multi-class classification, the random forest method has been widely applied in various scientific fields, such as metabolomic or microarray data.^[9-11] However, it has been less employed on cardiac data for the prediction of HF.

In the present study, a prediction model was constructed by combining four cardiac biomarkers, namely, CK-MB, BNP, Gal-3, and sST2, using the random forest algorithm. The importance of individual cardiac biomarkers for HF prediction was identified, and the predictive performance of the combination of the four biomarkers was analyzed. The present results suggest that the integration of CK-MB, BNP, Gal-3, and sST2 exhibited considerable improvement in assessing the risk of HF, when compared with individual biomarkers.

Methods

Ethical approval

The Ethics Committee of Beijing Anzhen Hospital approved the study protocol, and all patients provided a written informed consent prior to participation.

Study design and participants

This study was conducted in China from June 2017 to December 2017. A total of 100 consecutive patients with HF (71 male and 29 female), who consulted the Cardiology Department of Anzhen Hospital, were enrolled in the study. Exclusion criteria included acute myocardial infarction, atrial fibrillation, pulmonary hypertension, acute pulmonary embolism, interstitial lung diseases, metabolic disorders, connective tissue diseases, malignancies, sepsis, and blood system diseases, use of medication within 6 months before enrollment. All the HF patients were diagnosed and treated in accordance with the criteria of the 2016 European Society of Cardiology (ESC) Guidelines for the diagnosis and treatment of acute and chronic heart failure.^[12] The New York Heart Association (NYHA) classification was used to grade the symptoms of patients with HF. Among these patients, 20 patients were classified as class I to II, while 80 patients were classified as class III to IV. In addition, another 113 age- and gender-matched healthy individuals were included, which served as controls.

Data collection

Data of the patients were collected from medical charts and clinical examinations. Demographic information (ie, age and gender) and health behaviors (ie, smoking status and alcohol consumption) were obtained by self-reporting during a face-to-face interview. A physical examination, including anthropometric measurements (height and body weight) and blood pressure measurements, was conducted by specifically trained staff. Body weight and height were assessed while the subjects stood barefoot in light clothing.

Laboratory examination

Blood samples were collected after overnight fasting, and immediately centrifuged for examination or stored at -70°C until assayed. Blood biochemical indicators, including fasting plasma glucose (FPG), triglyceride (TG), total cholesterol (TC), low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), C-reactive protein (CRP), serum creatinine (SCr), alanine aminotransferase (ALT), and aspartate aminotransferase (AST), were measured.

Plasma BNP, Gal-3, and CK-MB were measured using a standard electrochemiluminescence immunoassay on an ARCHITECT I2000SR analyzer (Abbott Laboratories, Abbott Park, IL, USA). The sST2 was assayed via high sensitivity sandwich monoclonal immunoassay (PresageTM ST2 assay; Critical Diagnostics, NY, USA).

Echocardiographic measurements

Standard echocardiography with Doppler was performed by the same ultrasound technician. All study subjects underwent echocardiography after admission. Left ventricular ejection fraction (LVEF) was quantitatively measured using the Simpson two-dimensional methodology. Left ventricular end diastolic dimension (LVEDD) and left ventricular end systolic dimension (LVESD) were measured according to the American Society of Echocardiography guidelines.^[13]

Definition of variables

Smoking was defined as a current cigarette consumption of an average of at least one cigarette daily for at least 1 year. Drinking was defined as a current consumption of an average of at least 50 g of alcohol daily for at least one year. BMI was calculated as weight in kilograms divided by height in meters squared (kg/m^2). The estimated glomerular filtration rate (eGFR) value was calculated using the Chinese-modified Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equation.

Statistical analysis

Statistical analyses were performed using the SAS (version 9.4; SAS Institute, Inc., Cary, NC, USA) and R statistical package.

For statistical description, all continuous variables were described as mean \pm standard deviation (SD), or medians

and percentiles (25th percentile and 75th percentile). Comparisons of continuous variables between groups were performed using *t* test for normally distributed data and rank sum test for data not normally distributed. In addition, discrete variables were reported as numbers (*n*) and percentages (%), and compared using Chi-squared test. The Satterthwaite method was used for unequal variances.

The Spearman rank correlation analysis was conducted between cardiac biomarkers and echocardiographic parameters. Correlations between 0 and 0.4, and 0.7, and 1 indicate a weak, moderate and strong correlation, respectively. A correlation coefficient >0.7 , within 0.4 to 0.7 and <0.4 indicated a strong, moderate, and weak correlation, respectively. Linear regression analysis was performed to assess the relationship of the echocardiographic findings with the serum cardiac biomarkers.

The receiver operating characteristic (ROC) curve was drawn to determine the diagnostic power of the cardiac biomarkers for the prediction of HF, in terms of sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV). The area under the curve (AUC) was calculated to determine the optimal cut-off value of an individual predictor of HF, at the point when AUC value was maximized.

Random forests of CK-MB, BNP, sST2, and Gal-3 were established for prediction via the *randomForest* package in R software, where the number of trees was specified as 500 to obtain stable results. The importance of individual variables was measured by the mean decrease in accuracy and Gini measures. Variables with a higher mean decrease in accuracy or Gini value were considered to be more important than those with lower values.

Statistical significance was accepted as a two-sided test with a *P* value <0.05 .

Results

Characteristics of participants

A total of 193 participants were included in the present study. Among these participants, 80 participants were patients with HF, while 113 participants were control subjects. The characteristics of these participants are presented in Table 1. Overall, the mean age of these patients was 53.7 ± 14.1 years old, and 61.7% of these patients were male. Patients with HF tended to be male. Furthermore, differences in age, height, weight and the BMI value between HF patients and controls were not statistically significant. The presence of HF was associated with a higher incidence of current smoking and drinking. The value of SBP was higher in patients with HF than that of control subjects. When compared with control subjects, patients with HF were more likely to have higher values of ALT, AST, SCr, and CRP, and lower values of eGFR, TC, LDL-C, and HDL-C.

Patients with HF had poor cardiac function and abnormal ventricular dimensions, as evidenced by the lower value of LVEF and higher values of LVEDD and LVESD, and in

comparison to their counterparts. For cardiac biomarkers, patients with HF exhibited significantly higher levels of CK-MB, BNP, Gal-3, and sST2, compared with control subjects (all, $P > 0.05$).

Correlation and regression analysis

Spearman rank correlation analysis revealed that BNP was moderately correlated with NYHA ($R=0.4047$, $P=0.0004$). However, BNP was poorly correlated with echocardiographic measures (LVEF, LVEDD, and LVESD; all, $R < 0.4$) [Table 2]. The other three cardiac biomarkers, CK-MB, sST2 and Gal-3, were also poorly correlated with NYHA, as well as echocardiographic parameters (all, $R < 0.4$), expecting a weak but significant correlation between Gal-3 and NYHA ($R=0.3617$, $P=0.0011$).

The linear regression analysis revealed no significant relationship between any cardiac biomarker mentioned and echocardiographic parameters measured [Table 3].

Predictive performance of individual biomarkers for HF

BNP exhibited a good independent predictive capacity for HF, with a sensitivity, specificity and AUC of 94.4%, 90.3%, and 0.956, at a cut-off value of 60 pg/mL. At a cut-off of 100 pg/mL, the specificity of BNP increased (95.6%), but its sensitivity decreased (86.1%), for predicting HF [Figure 1].

However, CK-MB, sST2, and Gal-3 exhibited a modest diagnostic performance for HF, with an AUC of 0.709, 0.711, and 0.777, respectively [Table 4].

Random forests

A multi-marker model was constructed by combining four cardiac biomarkers (CK-MB, BNP, Gal-3, and sST2) using the random forest method. In order to minimize the possible variance of the present results, the random forest algorithm was run multiple times to obtain the average of the predictions. The results of the random forests are presented in Figure 2.

BNP was the most important variable, as evidenced by a significantly higher mean decrease in accuracy and Gini, when compared with sST2, Gal-3, and CK-MB. Overall, there was a general increase in predictive performance using the multi-marker model, when compared with individual cardiac biomarkers, with a sensitivity, specificity, PPV, and NPV of 91.5%, 96.7%, 97.0%, and 90.8%.

Discussion

The present study analyzed a multi-marker approach via the random forest algorithm for the prediction of HF. The random forests, which were implemented with 500 trees when training, revealed that BNP was the most important cardiac biomarker for HF prediction, as assessed by the mean decrease in accuracy and Gini measures. This was consistent with the ROC results for individual cardiac

Table 1: Baseline characteristics of 80 patients with HF and 113 age- and gender-matched healthy controls.

Variables	Total	Control group	HF group	Statistics	P value
Clinical					
Male, <i>n</i> (%)	119 (61.7)	61 (54.0)	58 (72.5)	6.79*	0.0091
Smoking, <i>n</i> (%)	66 (34.9)	27 (24.6)	39 (49.4)	12.47*	0.0004
Drinking, <i>n</i> (%)	47 (24.9)	21 (19.1)	26 (32.9)	4.70*	0.0302
Age, years	54.0 (44.0, 62.0)	53.0 (44.0, 62.0)	57.0 (43.5, 63.5)	0.69†	0.4921
Height, cm	166.0 (159.0, 172.0)	165.0 (158.0, 171.0)	170.0 (160.0, 174.0)	1.94†	0.0521
Weight, kg	67.0 (58.0, 79.0)	66.5 (58.0, 77.5)	70.0 (59.0, 83.0)	1.09†	0.2747
BMI, kg/m ²	24.8 (22.4, 27.4)	24.6 (22.5, 27.2)	25.3 (22.0, 28.7)	0.72†	0.5751
SBP, mmHg	120.0 (109.0, 138.0)	122.0 (115.0, 141.0)	111.0 (100.0, 130.0)	-3.55†	0.0004
DBP, mmHg	73.0 (65.0, 81.0)	72.0 (65.0, 81.0)	75.0 (62.0, 84.0)	0.86†	0.3891
HR, beats/min	79.0±17.3	64.8±6.6	81.7±17.4	-6.51‡	<0.0001
Echocardiography					
LVEF, %	60.0 (32.0, 66.0)	65.0 (62.0, 68.0)	30.0 (25.0, 41.5)	-10.05†	<0.0001
LVEDD, mm	50.0 (46.0, 62.0)	46.0 (45.0, 49.0)	64.0 (57.5, 70.5)	9.46†	<0.0001
LVESD, mm	34.0 (29.0, 51.0)	30.0 (28.0, 32.0)	54.0 (45.0, 60.0)	9.64†	<0.0001
Laboratory					
CK-MB, ng/mL	1.6 (1.0, 2.4)	1.2 (0.9, 1.6)	1.9 (1.4, 2.9)	-4.25†	<0.0001
BNP, pg/mL	44.0 (18.0, 384.0)	24.0 (13.0, 41.0)	742.5 (256.0, 1390.0)	10.45†	<0.0001
sST2, ng/mL	16.87 (13.26, 22.26)	15.21 (12.69, 18.35)	21.81 (14.88, 27.42)	-4.40†	<0.0001
Gal-3, ng/mL	11.50 (8.80, 15.80)	10.00 (8.00, 12.80)	15.75 (11.10, 19.75)	-7.19†	<0.0001
ALT, U/L	20.0 (15.0, 34.0)	18.0 (14.0, 24.0)	29.5 (19.0, 50.5)	4.90†	<0.0001
AST, U/L	24.00 (19.00, 31.50)	20.00 (18.00, 24.50)	28.50 (24.00, 39.00)	-6.15†	<0.0001
SCr, μmol/L	69.60 (56.80, 80.60)	64.00 (54.30, 73.60)	78.30 (68.00, 100.05)	-6.77†	<0.0001
eGFR, mL·min ⁻¹ ·1.73 m ⁻²	100.08 (83.23, 108.88)	103.21 (95.93, 110.01)	87.17 (69.88, 102.26)	-5.24†	<0.0001
FPG, mmol/L	5.48 (5.09, 6.18)	5.46 (5.16, 5.88)	5.55 (4.95, 6.56)	-0.08†	0.5276
TG, mmol/L	1.19 (0.78, 1.68)	1.22 (0.84, 1.86)	1.14 (0.72, 1.59)	-1.19†	0.2324
TC, mmol/L	4.58 (3.93, 5.34)	4.78 (4.38, 5.63)	3.99 (3.27, 4.79)	-5.32†	<0.0001
LDL-C, mmol/L	2.76 (2.16, 3.37)	2.89 (2.40, 3.46)	2.55 (1.67, 3.27)	-3.04†	0.0024
HDL-C, mmol/L	1.14 (0.95, 1.44)	1.33 (1.10, 1.55)	0.93 (0.73, 1.12)	-8.05†	<0.0001
CRP, mg/dL	1.00 (0.52, 2.46)	0.84 (0.42, 1.58)	1.73 (0.75, 4.69)	4.63†	<0.0001
Hemoglobin, g/L	141.00 (128.00, 154.00)	141.00 (131.00, 152.00)	140.50 (124.00, 157.50)	-0.04†	0.9656

Data are expressed as *n* (%), mean±standard deviation, or median (Q1, Q3). * χ^2 value by Chi-square test. †Z value by Rank sum test. ‡*t* value by Satterthwaite method. ALT: Alanine aminotransferase; AST: Aspartate aminotransferase; BMI: Body mass index; BNP: Brain natriuretic peptide; CK-MB: Creatinine kinase-MB isoenzyme; CRP: C-reactive protein; DBP: Diastolic blood pressure; eGFR: Estimated glomerular filtration rate; FPG: Fasting plasma glucose; Gal-3: galectin-3; HDL-C: High-density lipoprotein cholesterol; HF: Heart failure; LDL-C: Low-density lipoprotein cholesterol; LVEF: Left ventricular ejection fraction; LVEDD: Left ventricular end-diastolic dimension; LVESD: Left ventricular end-systolic dimension; sST2: Soluble suppression of tumorigenicity 2; SBP: Systolic blood pressure; SCr: Serum creatinine; TG: Triglyceride; TC: Total cholesterol.

Table 2: Result of Spearman rank correlation analysis between cardiac biomarkers and echocardiographic parameters.

Markers	NYHA		LVEDD		LVESD		LVEF	
	Correlation coefficient	P value	Correlation coefficient	P value	Correlation coefficient	P value	Correlation coefficient	P value
CK-MB	0.0328	0.7754	0.0433	0.7777	0.0582	0.7076	0.1410	0.2408
BNP	0.4047	0.0004	0.1845	0.2482	0.2285	0.1561	-0.0743	0.5564
sST2	0.0548	0.6315	0.0482	0.7503	0.0976	0.5237	-0.2278	0.0543
Gal-3	0.3617	0.0011	0.1659	0.2705	0.2277	0.1325	-0.1704	0.1523

BNP: B-type natriuretic peptide; CK-MB: Creatine kinase MB isoenzyme; Gal-3: Galectin-3; LVEDD: Left ventricular end diastolic dimension; LVESD: Left ventricular end systolic dimension; LVEF: Left ventricular ejection fraction; NYHA: New York Heart Association; sST2: Soluble suppression of tumorigenicity-2.

predictors, which exhibited good sensitivity and specificity for BNP. In addition, the random forest algorithm revealed that the combination of the four cardiac biomarkers (CK-MB, BNP, Gal-3, and sST2) achieved a general increase in prediction accuracy for HF, suggesting its potential application for the clinical assessment of patients with HF.

BNP is a neurohormone synthesized in the cardiac ventricles, and its release is directly proportional to ventricular expansion and pressure overload.^[14] Furthermore, BNP concentrations are higher in patients with more advanced NYHA classes,^[15] demonstrating a correlation between plasma BNP concentration and the NYHA

Table 3: Result of linear regression analysis between cardiac biomarkers and imaging parameters.

Markers	LVEDD			LVESD			LVEF		
	Regression coefficient	t value	P value	Regression coefficient	t value	P value	Regression coefficient	t value	P value
CK-MB	0.0058	0.00	0.9996	1.8229	0.20	0.8462	1.0036	1.60	0.1145
BNP	0.0136	1.01	0.3206	0.0156	1.28	0.2090	0.0000	0.01	0.9882
sST2	1.2331	0.86	0.3939	1.3065	1.01	0.3173	-0.1617	-1.65	0.1043
Gal-3	1.3207	0.65	0.5196	2.0836	1.14	0.2594	-0.1585	-0.85	0.4007

BNP: B-type natriuretic peptide; CK-MB: Creatine kinase MB isoenzyme; Gal-3: Galectin-3; LVEDD: Left ventricular end diastolic dimension; LVESD: Left ventricular end systolic dimension; LVEF: Left ventricular ejection fraction; sST2: Soluble suppression of tumorigenicity-2.

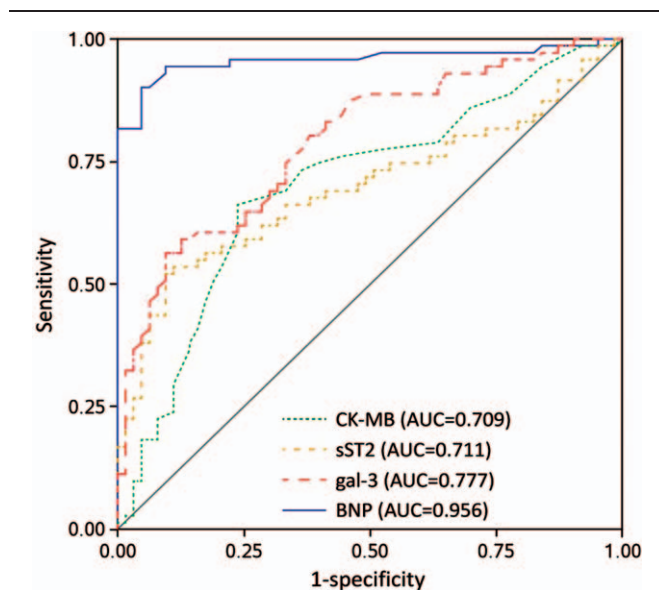


Figure 1: Predictive performance of individual biomarkers for heart failure. BNP exhibited a good independent predictive capacity for HF, with an AUC of 0.956. CK-MB, sST2, and Gal-3 exhibited a modest diagnostic performance for HF, with an AUC of 0.709, 0.711, and 0.777, respectively.

classification.^[16-18] Consistently, it was found that among the four cardiac biomarkers, only plasma BNP concentration moderately correlated with the NYHA class. The combined assessment of cardiac biomarkers together with echocardiography provides more a powerful risk stratification for patients with HF across all stages.^[19] The relationship between BNP and echocardiographic measurements remain uncertain with mixed results, depending upon the study population. In some studies, plasma BNP concentration correlates moderately or strongly with echocardiographic indexes.^[20-21] However, other studies have revealed that echocardiographic indexes and circulating cardiac indicators (BNP and Gal-3) were poorly or even not correlated in patients with HF,^[22-23] which were consistent with the present results. In addition, linear regression analysis indicated no relationship among any of the cardiac biomarkers assessed and echocardiographic parameters measured.

The results for single plasma cardiac biomarkers, in terms of predictive capacity, have been conflicting. The Breathing Not Properly Study measured plasma BNP concentrations in 1586 patients with acute dyspnea, and found that BNP was the best single predictor for predicting the final

diagnosis of HF among demographic indicators, imaging parameters and laboratory variables.^[18] Furthermore, a study revealed the superior biochemically diagnostic capacity of BNP over Gal-3 and sST2 for HF.^[24] Consistently, a good predictive performance of BNP for HF diagnosis was found, with a sensitivity, specificity and AUC of 94.4%, 90.3%, and 0.956, respectively, at the cut-off of 60 pg/mL. However, CK-MB, sST2, and Gal-3 only exhibited modest discriminability for HF patients. Despite acceptable sensitivity, sST2 has been considered to lack of specificity for the prediction of HF, limiting its independent application as a diagnostic tool alone.^[25] In a study of 599 patients with acute dyspnea, Gal-3 exhibited a moderate diagnostic performance for acute HF with an AUC of 0.72 *vs.* an AUC of 0.94 for BNP.^[26] In contrast to BNP, both galectin-3 and sST2 were questioned to be probably not useful as an aid for the diagnosis of HF.^[24] Interestingly, the head-to-head comparison of sST2 *vs.* Gal-3 in chronic HF demonstrated the superiority of ST2 over Gal-3 in risk stratification.^[27] The present findings demonstrated an opposite result, in which the AUC of Gal-3 was slightly greater than that of sST2 for predicting HF.

The random forest method represents as a useful tool to identify the most important predictors from a collection of variables by calculating several measures. One of the measures was mean decrease accuracy, which ranks the importance of the predictor based on the decrease in prediction accuracy, when the values of the variable are randomly permuted. The Gini index is another measure of the prediction powers of variables through the sum of all decreases in Gini impurity, which measures how often an element would be incorrectly labeled if randomly labeled according to the distribution of labels in the subset. Thus, a greater mean decrease in accuracy or Gini represents that a predictor feature plays a more important role in partitioning the data at a node of the decision tree. Ward *et al* used random forests to identify and validate important predictors of mortality among patients with systemic lupus erythematosus, achieving an AUC of 0.94.^[28] In the present study, the random forest test was performed for four predictors, and BNP displayed a greater mean decrease accuracy and Gini, compared with sST2, Gal-3, and CK-MB, highlighting it as the most important individual predictor among the variables examined for predicting HF.

Since no single factor reliably predicts HF, attempts have been made to improve risk prediction by combining a

Table 4: Prediction properties of single cardiac biomarkers.

Predictors		Cut-off point	AUC	Sensitivity	Specificity	PPV	NPV
BNP (pg/mL)	(Model)	60.0	0.956	0.944	0.903	0.861	0.962
	(Reference)	100.0	–	0.861	0.956	0.925	0.915
CK-MB (ng/mL)	(Model)	1.7	0.709	0.658	0.762	0.776	0.640
sST2 (ng/mL)	(Model)	21.6	0.711	0.525	0.876	0.750	0.723
Gal-3 (ng/mL)	(Model)	15.0	0.777	0.575	0.885	0.780	0.746

AUC: Area under the curve; BNP: B-type natriuretic peptide; CK-MB: Creatine kinase MB isoenzyme; Gal-3: Galectin-3; NPV: Negative predictive value; PPV: Positive predictive value; sST2: Soluble suppression of tumorigenicity-2.

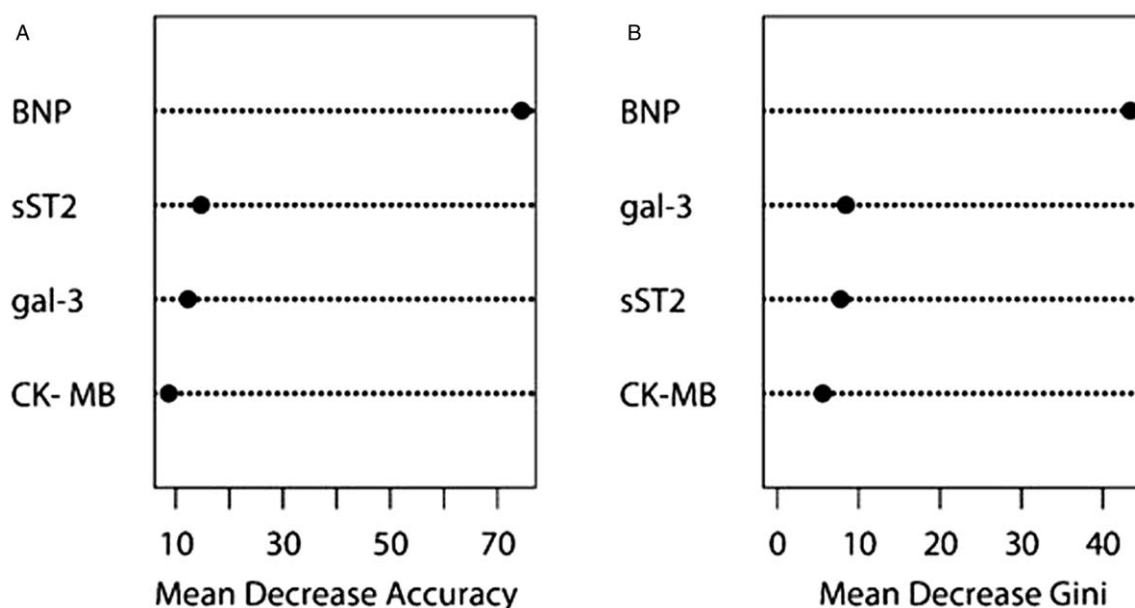


Figure 2: The importance of variables in the random forest algorithm. A multi-marker model was constructed by combining four cardiac biomarkers (CK-MB, BNP, Gal-3, and sST2) using the random forest algorithm. BNP was the most important variable, as evidenced by its significantly higher mean decrease accuracy (A) and mean decrease in Gini (B), when compared with sST2, Gal-3, and CK-MB.

diverse panel of biomarkers, allowing for the integration of various pathophysiological aspects of the disease process, such as myocardial injury, vascular load, neurohormonal activation and inflammation.^[29] The combination of ST2 and BNP offers a modest improvement in the risk stratification of chronic HF patients.^[30] Combining plasma Gal-3 and BNP increased the diagnostic and prognostic value over either of biomarkers alone.^[31-32] In a multi-center cohort of 1513 ambulatory HF patients, a multi-marker panel that comprised of seven circulating biomarkers was reflective of diverse biological pathways, and was proven to be a strong predictor that outperformed the Seattle Heart Failure Model (SHFM) score, and substantially improved the prediction of chronic HF.^[33] These findings support the potential use of the multi-marker approach for improving the predictive power for HF. In the present study, although BNP had a good predictive performance for HF, the improvement in accuracy was observed by the addition of CK-MB, Gal-3, and sST2 to BNP for the biochemical diagnosis of HF, suggesting the meaningfulness of a multi-marker strategy for HF diagnosis.

Although the concept of a multi-marker tool was endorsed, determining how to identify an optimal panel of biomarkers for assessing HF is a formidable task. In order to classify patients and identify predictors, random forests were employed, which is a novel method that develops numerous decision trees, by which the accuracy of the predictors were tested. The random forest classifier operates by building a set of decision trees, and at each node in the trees, a random subset of the predictor variables are randomly selected and considered as split candidates. The dataset was repeatedly divided into subtrees, assessing predictor variables by importance on a basis of the change in the classification error affected by its presence or absence in the subset. Furthermore, the random forests also combine the predictions of multiple decision trees, improving the power of the algorithm.^[34] In the present study, the implementation of a random forest algorithm revealed that the combination of BNP, CK-MB, Gal-3, and sST2 achieved the greatest accuracy for HF prediction, with the sensitivity, specificity, PPV and NPV all exceeding 90%. This outperformed the models for individual cardiac biomarker assessment, suggesting an

improvement in overall accuracy via a multi-marker strategy.

The present study has several limitations. First, this was a preliminary observational study with relatively small sample size that was performed in a single center, and the relevance of these results among patients in other parts of the world was not validated. Second, given the potential confounding and selection bias by indication, the cautious interpretation of these data is needed. Third, the present study focused on four cardiac biomarkers (BNP, CK-MB, sST2, and Gal-3), and some other biomarkers reflective of other aspects of the underlying pathophysiology of HF, such as renal salt and water retention, and oxidative stress, were not included. These limitations could be overcome through future large-scale, prospective studies with multiple attempts of combining a diverse panel of biomarkers. Also, for four biomarkers, especially galectin-3 and ST2 which are mentioned by the 2013 American College of Cardiology/American Heart Association guideline for the management of heart failure,^[35] their prognostic value (eg, for hospitalization and mortality) in patients with HF deserves further investigation.

In conclusion, the present study provided a framework for the exploration of the random forest algorithm in the prediction of HF. The present findings support that BNP has a higher accuracy for the diagnosis of HF, compared with CK-MB, galectin-3, and sST2. In addition, CK-MB, Gal-3, and sST2 to BNP led to a substantial improvement in the biochemical diagnosis of HF. The random forest algorithm provides a robust method to assess the accuracy and importance of predictor variables. In addition, its potential to validate the usefulness of multiple biomarkers for HF diagnosis needs further investigation.

Funding

This work was supported by grants from the National Natural Science Foundation of China (No. 81770353) and the Abbott China research fund (ADD-2017).

Conflicts of interest

This study received Abbott China research fund.

References

- Roger VL. Epidemiology of heart failure. *Circ Res* 2013;113:646–659. doi: 10.1093/med/9780198784906.003.0401.
- Gaggin HK, Januzzi JL Jr. Biomarkers and diagnostics in heart failure. *Biochim Biophys Acta* 2013;1832:2442–2450. doi: 10.1016/j.bbdis.2012.12.014.
- Apple FS. Creatine kinase isoforms and myoglobin: early detection of myocardial infarction and reperfusion. *Coron Artery Dis* 1999;10:75–79. doi: 10.1097/00019501-199910001-00003.
- Kakkar R, Lee RT. The IL-33/ST2 pathway: therapeutic target and novel biomarker. *Nat Rev Drug Discov* 2008;7:827–840. doi: 10.1038/nrd2660.
- Rehman SU, Mueller T, Januzzi JL Jr. Characteristics of the novel interleukin family biomarker ST2 in patients with acute heart failure. *J Am Coll Cardiol* 2008;52:1458–1465. doi: 10.1016/j.jacc.2008.07.042.
- Ho JE, Liu C, Lyass A, Courchesne P, Pencina MJ, Vasan RS, *et al.* Galectin-3, a marker of cardiac fibrosis, predicts incident heart failure in the community. *J Am Coll Cardiol* 2012;60:1249–1256. doi: 10.1016/j.jacc.2012.04.053.
- Braunwald E. Biomarkers in heart failure. *N Engl J Med* 2008;358:2148–2159. doi: 10.1056/NEJMra0800239.
- Leo B, Adele C. Random forests, http://statwww.berkeley.edu/users/breiman/RandomForests/cc_home.htm, University of California, Berkeley, CA, USA.
- Shi T, Seligson D, Belldegrun AS, Palotie A, Horvath S. Tumor classification by tissue microarray profiling: random forest clustering applied to renal cell carcinoma. *Mod Pathol* 2005;18:547–557. doi: 10.1038/modpathol.3800322.
- Chen T, Cao Y, Zhang Y, Liu J, Bao Y, Wang C, *et al.* Random forest in clinical metabolomics for phenotypic discrimination and biomarker selection. *Evid Based Complement Alternat Med* 2013;2013:298183. doi: 10.1155/2013/298183.
- Chen X, Ishwaran H. Random forests for genomic data analysis. *Genomics* 2012;99:323–329. doi: 10.1016/j.ygeno.2012.04.003.
- Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JGF, Coats AJS, *et al.* 2016 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure. *Rev Esp Cardiol (Engl Ed)* 2016;69:1167. doi: 10.1016/j.rec.2016.11.005.
- Lang RM, Badano LP, Mor-Avi V, Afilalo J, Armstrong A, Ernande L, *et al.* Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *Eur Heart J Cardiovasc Imaging* 2015;16:233–270. doi: 10.1093/ehjci/jev014.
- Maisel A. B-type natriuretic peptide levels: diagnostic and therapeutic potential. *Cardiovasc Toxicol* 2001;1:159–164. doi: 10.1385/ct:1:2:159.
- Palazzuoli A, Gallotta M, Quatrini I, Nuti R. Natriuretic peptides (BNP and NT-proBNP): measurement and relevance in heart failure. *Vasc Health Risk Manag* 2010;6:411–418. doi: 10.2147/vhrm.s5789.
- Eindhoven JA, van den Bosch AE, Jansen PR, Boersma E, Roos-Hesselink JW. The usefulness of brain natriuretic peptide in complex congenital heart disease: a systematic review. *J Am Coll Cardiol* 2012;60:2140–2149. doi: 10.1016/j.jacc.2012.02.092.
- Kemperman H, van den Berg M, Kirkels H, de Jonge N. B-type natriuretic peptide (BNP) and N-terminal proBNP in patients with end-stage heart failure supported by a left ventricular assist device. *Clin Chem* 2004;50:1670–1672. doi: 10.1373/clinchem.2003.030510.
- Maisel AS, Krishnaswamy P, Nowak RM, McCord J, Hollander JE, Duc P, *et al.* Rapid measurement of B-type natriuretic peptide in the emergency diagnosis of heart failure. *N Engl J Med* 2002;347:161–167. doi: 10.1056/NEJMoa020233.
- Troughton RW, Richards AM. B-type natriuretic peptides and echocardiographic measures of cardiac structure and function. *JACC Cardiovasc Imaging* 2009;2:216–225. doi: 10.1016/j.jcmg.2008.12.006.
- Troughton RW, Prior DL, Pereira JJ, Martin M, Fogarty A, Morehead A, *et al.* Plasma B-type natriuretic peptide levels in systolic heart failure: importance of left ventricular diastolic function and right ventricular systolic function. *J Am Coll Cardiol* 2004;43:416–422. doi: 10.1016/j.jacc.2003.08.046.
- Kuster GM, Tanner H, Printzen G, Suter TM, Mohacsi P, Hess OM. B-type natriuretic peptide for diagnosis and treatment of congestive heart failure. *Swiss Med Wkly* 2002;132:623–628. doi: 2002/43/smw-10081.
- Fornwalt BK, Sprague WW, BeDell P, Suever JD, Gerritse B, Merlino JD, *et al.* Agreement is poor among current criteria used to define response to cardiac resynchronization therapy. *Circulation* 2010;121:1985–1991. doi: 10.1161/CIRCULATIONAHA.109.910778.
- Hrynchyshyn N, Jourdain P, Desnos M, Diebold B, Funck F. Galectin-3: a new biomarker for the diagnosis, analysis and prognosis of acute and chronic heart failure. *Arch Cardiovasc Dis* 2013;106:541–546. doi: 10.1016/j.acvd.2013.06.054.
- Mueller T, Gegenhuber A, Leitner I, Poelz W, Haltmayer M, Dieplinger B. Diagnostic and prognostic accuracy of galectin-3 and soluble ST2 for acute heart failure. *Clin Chim Acta* 2016;463:158–164. doi: 10.1016/j.cca.2016.10.034.
- Dieplinger B, Mueller T. Soluble ST2 in heart failure. *Clin Chim Acta* 2015;443:57–70. doi: 10.1016/j.cca.2014.09.021.

26. van Kimmenade RR, Januzzi JL Jr, Ellinor PT, Sharma UC, Bakker JA, Low AF, *et al.* Utility of amino-terminal pro-brain natriuretic peptide, galectin-3, and apelin for the evaluation of patients with acute heart failure. *J Am Coll Cardiol* 2006;48:1217–1224. doi: 10.1016/j.jacc.2006.03.061.
27. Bayes-Genis A, de Antonio M, Vila J, Peñafiel J, Galán A, Barallat J, *et al.* Head-to-head comparison of 2 myocardial fibrosis biomarkers for long-term heart failure risk stratification: ST2 versus galectin-3. *J Am Coll Cardiol* 2014;63:158–166. doi: 10.1016/j.jacc.2013.07.087.
28. Ward MM, Pajevic S, Dreyfuss J, Malley JD. Short-term prediction of mortality in patients with systemic lupus erythematosus: classification of outcomes using random forests. *Arthritis Rheum* 2006;55:74–80. doi: 10.1002/art.21695.
29. Bayes-Genis A, Ordonez-Llanos J. Multiple biomarker strategies for risk stratification in heart failure. *Clin Chim Acta* 2015;443:120–125. doi: 10.1016/j.cca.2014.10.023.
30. Ky B, French B, McCloskey K, Rame JE, McIntosh E, Shahi P, *et al.* High-sensitivity ST2 for prediction of adverse outcomes in chronic heart failure. *Circ Heart Fail* 2011;4:180–187. doi: 10.1161/CIRCHEARTFAILURE.110.958223.
31. de Boer RA, Lok DJ, Jaarsma T, van der Meer P, Voors AA, Hillege HL, *et al.* Predictive value of plasma galectin-3 levels in heart failure with reduced and preserved ejection fraction. *Ann Med* 2011;43:60–68. doi: 10.3109/07853890.2010.538080.
32. Shah RV, Chen-Tournoux AA, Picard MH, van Kimmenade RR, Januzzi JL. Galectin-3, cardiac structure and function, and long-term mortality in patients with acutely decompensated heart failure. *Eur J Heart Fail* 2010;12:826–832. doi: 10.1093/eurjhf/hfq091.
33. Ky B, French B, Levy WC, Sweitzer NK, Fang JC, Wu AH, *et al.* Multiple biomarkers for risk prediction in chronic heart failure. *Circ Heart Fail* 2012;5:183–190. doi: 10.1161/CIRCHEARTFAILURE.111.965020.
34. Sylvester EVA, Bentzen P, Bradbury IR, Clément M, Pearce J, Horne J, *et al.* Applications of random forest feature selection for fine-scale genetic population assignment. *Evol Appl* 2017;11:153–165. doi: 10.1111/eva.12524.
35. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE Jr, Drazner MH, *et al.* 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on practice guidelines. *Circulation* 2013;128:e240–327. doi: 10.1016/j.jacc.2013.05.019.

How to cite this article: Yuan H, Fan XS, Jin Y, He JX, Gui Y, Song LY, Song Y, Sun Q, Chen W. Development of heart failure risk prediction models based on a multi-marker approach using random forest algorithms. *Chin Med J* 2019;132:819–826. doi: 10.1097/CM9.0000000000000149