

## Research Article

# Hao Li, Yuan Cui, Jin Tian, Hong Yang, Qing Zhang, Ke Wang, Qinghua Han\*, Yanbo Zhang\* A novel age-biomarker-clinical history prognostic index for heart failure with reduced left ventricular ejection fraction

<https://doi.org/10.1515/med-2020-0209>

received January 27, 2020; accepted June 10, 2020

## Abstract

**Purpose** – A model for predicting the prognosis of patients with heart failure with reduced left ventricular ejection fraction (HFrEF) is currently not available. This study aimed to develop an age-biomarker-clinical history prognostic index (ABC-PI) and validate it for the assessment of individual prognosis.

**Patients and methods** – A total of 5,974 HFrEF patients were enrolled and 1,529 were included in this study after excluding missing values and loss to follow-up. Variables that significantly contributed to prediction of all-cause mortality were assessed by Cox regression and latent trait analysis (LTA) was used to validate discrimination of variables.

**Results** – After Cox regression, the following seven most significant variables were selected: age, N-terminal pro-B-type natriuretic peptide, renal dysfunction, left ventricular mass index, percutaneous coronary intervention, atrial fibrillation, and New York Heart Association (C-index:  $0.801 \pm 0.013$ ). After verification by LTA, discrimination of these seven variables was proven. A nomogram was used to

form the ABC-PI, and then the total score was set to 100 points. A lower score indicated a higher risk. After verification, the 3-year mortality rate was 34.7% in the high-risk group and only 2.6% in the low-risk group.

**Conclusion** – Our novel ABC-PI shows a good performance and does not require re-input in the original model. The ABC-PI can be used to effectively and practically predict the prognosis of HFrEF patients.

**Keywords:** HFrEF, latent trait analysis, nomogram, NT-proBNP

## 1 Introduction

Approximately 26 million people suffer from heart failure worldwide, and it has become a global public health problem [1]. Heart failure is the most common cardiovascular cause of hospitalization in patients older than 60 years [2]. Heart failure has a high prevalence and mortality rate, and it can severely impair physical function and quality of life [3,4]. Heart failure with reduced left ventricular ejection fraction (HFrEF) is a complex condition. Patients with HFrEF may belong to different subpopulations and be associated with different risks of death. Therefore, creating an HFrEF prognosis model for determining the risk of death of these patients is desirable.

Previous studies on cardiovascular disease often focused on risk score systems to identify the risk of certain events rather than prognosis, such as using scores to predict the risk of bleeding in patients with atrial fibrillation [5]. Furthermore, in many classic risk scoring systems, biomarker variables, such as Framingham risk functions, are not included [6]. Measurement of biomarkers has become a routine examination during hospitalization, and it has a strong predictive power. Therefore, use of biomarkers should be considered when creating a prognostic index (PI). There are many studies exploring the biomarkers of cardiovascular disease [7,8]. A common effective biomarker in patients with HFrEF is N-terminal pro-B-type natriuretic peptide (NT-proBNP) [9]. Currently, the most commonly used indicator in patients with HFrEF is the New York Heart Association (NYHA) functional classification system, which

---

\* **Corresponding author: Qinghua Han**, Department of Cardiology, First Hospital of Shanxi Medical University, 85 South Xinjian Road, Taiyuan, Shanxi Province 030001, People's Republic of China, e-mail: hanqinghua11@126.com, tel: +86-139-3546-2672

\* **Corresponding author: Yanbo Zhang**, Department of Health Statistics, Shanxi Provincial Key Laboratory of Major Diseases Risk Assessment, School of Public Health, Shanxi Medical University, 56 South Xinjian Road, Taiyuan, Shanxi Province 030001, People's Republic of China, e-mail: sxmuzyb@126.com, tel: +86-155-0368-5368

**Hao Li, Yuan Cui, Hong Yang, Qing Zhang, Ke Wang:** Department of Health Statistics, Shanxi Provincial Key Laboratory of Major Diseases Risk Assessment, School of Public Health, Shanxi Medical University, 56 South Xinjian Road, Taiyuan, Shanxi Province 030001, People's Republic of China

**Jin Tian:** Department of Cardiology, First Hospital of Shanxi Medical University, 85 South Xinjian Road, Taiyuan, Shanxi Province 030001, People's Republic of China

ORCID: Hao Li 0000-0002-8147-7542; Yanbo Zhang 0000-0002-9461-4163

was developed by the American College of Cardiology Foundation and the American Heart Association [10]. In this system, the NYHA classification is based on the patient's symptoms and the degree to which the condition affects their daily activities. This classification also needs to be considered as an important variable for inclusion in a PI. Additionally, there are some other common predictors, such as atrial fibrillation [11].

Latent trait analysis (LTA), which is a form of latent structure analysis, was first proposed in 1968 [12]. LTA is used for measurement characteristics in different subpopulations, with educational testing and mental testing in particular. This method has been proven to be effective [13].

This study aimed to develop a new age-biomarker-clinical history prognostic index (ABC-PI) and to validate it for predicting the prognosis of patients with HFrEF. Discrimination of the ABC-PI was tested by LTA. This study established a PI using a nomogram that was expected not to require re-input of the original model to achieve the purpose of improving practicability and effectiveness.

## 2 Methods

### 2.1 Study population

Patients with HFrEF were enrolled at the First Hospital of Shanxi Medical University and Shanxi Cardiovascular Hospital. The cohort was created in January 2014, and the cutoff date for analyses was August 2019. The inclusion criteria were an age of  $\geq 18$  years; diagnosis of HFrEF (presence of a basic cardiovascular disease, such as coronary heart disease and hypertension, with typical symptoms of chronic heart failure, such as paroxysmal dyspnea, tiredness, palpitation, pulmonary rales, and pleural effusion); an ejection fraction of  $< 40\%$  or  $40\text{--}50\%$  with structural heart disease or diastolic dysfunction; and NYHA grade II, III, or IV. The exclusion criteria were the presence of a mental disease or another serious disease (e.g., malignant tumors). Patients had data recorded at baseline and were followed up via cell phone and annual visits for 3 years. After each patient was discharged from the hospital, they were followed up via cell phone (the phone number was recorded on their first admission to the hospital) to confirm survival status, and in cases of death, the exact date of death was recorded.

A total of 5,974 patients were initially included. A total of 1,529 patients were finally included in the study after excluding incomplete variable data and loss to follow-up.

The patients were divided into two groups, including one for development of the PI ( $n = 878$ ) and another for verification of the PI ( $n = 651$ ).

### 2.2 HFrEF-electronic Case Report Form

An HFrEF-electronic Case Report Form (HFrEF-eCRF) was established to collect baseline information during hospitalization, which include the following: (1) basic information, such as demographic characteristics, admission time, discharge time, and clinical diagnosis; (2) comorbidities and history of allergies; (3) results of a standard physical examination; (4) laboratory test results; (5) results of imaging, including electrocardiography and echocardiography; and (6) medication and surgery information. A total of 222 variables were included in this study.

The HFrEF-eCRF was developed on the basis of clinical guidelines and clinicians' opinions. Entry and follow-up personnel were trained, and data quality was controlled using dual entry. The study was reviewed and approved by the Medical Ethics Committee of Shanxi Medical University, China (No. 2013099), and written informed consent was obtained from all participants.

### 2.3 Statistical analysis

Mean values and standard deviations were calculated for continuous data, and numbers and percentages were calculated for categorical data. In the development stage, univariable and multivariate logistic regression was used to initially screen the variables. To improve predictive efficiency, the selected continuous variables were transformed into categorical variables via receiver operating characteristic (ROC) curve segmentation and re-introduced into model validation. A log-rank procedure was then used to verify the variables and the validated variables were included in Cox stepwise regression analysis.

LTA was used to test discrimination of selected variables. LTA is also called item response theory in the field of psychometrics. This method establishes a functional relationship between latent categories and external variables. To date, there are at least 20 types of latent trait models. According to different data, different models can be used to estimate the parameters. In our study, the most common Rasch model and Ltm were used. The Rasch model is a special case of a one-dimensional latent trait model and the Ltm is a two-parameter logistic model [14]. Parameter

**Table 1:** Patients' characteristics

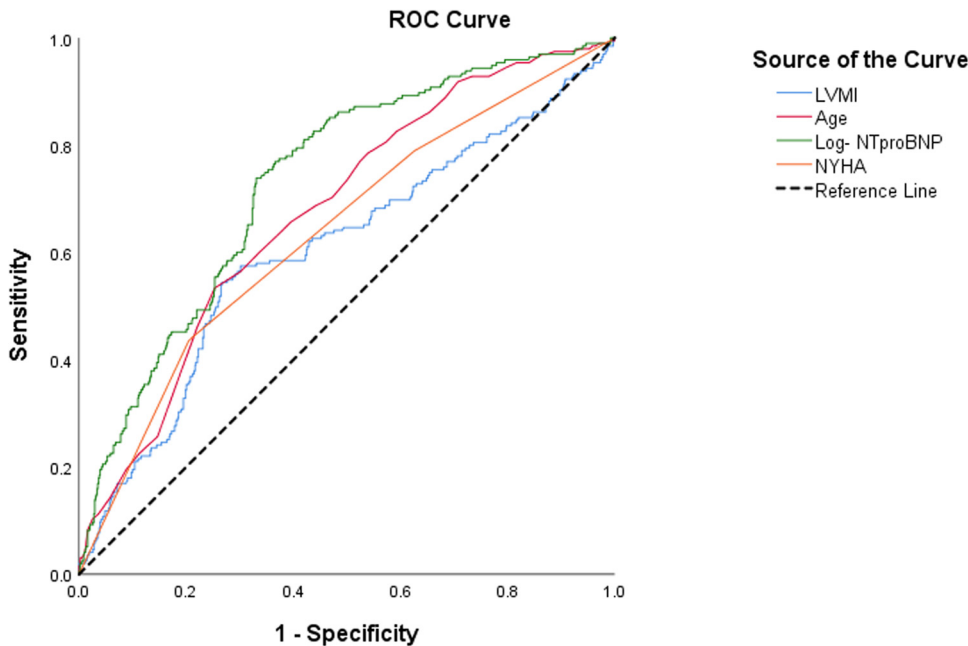
Characteristic	N	Survival (n = 683)	Death (n = 195)	OR (95% CI)	P
Basic data					
Age (years)	878	67.91 ± 10.97	74.31 ± 9.92	1.065 (1.046–1.084)	<0.001
Sex					
Male	585	468 (80.0%)	117 (20.0%)		
Female	293	215 (73.4%)	78 (26.6%)	1.451 (1.044–2.016)	0.026
NYHA grade					
II	296	255 (86.1%)	41 (13.9%)		
III	357	288 (80.7%)	69 (19.3%)	1.490 (0.977–2.272)	0.064
IV	225	140 (62.2%)	85 (37.8%)	3.766 (2.466–5.781)	<0.001
Heart rate (per minute)	878	75.01 ± 16.07	79.95 ± 16.87	1.017 (1.008–1.027)	<0.001
Body mass index (kg/m <sup>2</sup> )	878	24.78 ± 3.49	23.49 ± 3.92	0.899 (0.856–0.943)	<0.001
Biochemical data					
Log NT-proBNP (ng/L)	878	3.08 ± 0.50	3.49 ± 0.45	5.978 (4.095–8.727)	<0.001
Hemoglobin (g/L)	878	136.61 ± 17.88	127.90 ± 21.98	0.976 (0.968–0.985)	<0.001
Albumin (g/L)	878	42.17 ± 4.88	39.87 ± 4.86	0.902 (0.869–0.935)	<0.001
γ-Glutamyl transpeptidase (U/L)	878	42.86 ± 54.47	52.84 ± 67.04	1.003 (1.000–1.005)	0.041
Creatinine (mmol/L)	878	90.83 ± 40.31	110.45 ± 66.32	1.008 (1.004–1.011)	<0.001
Other findings					
PCI					
–	629	468 (74.4%)	161 (25.6%)		
+	249	215 (86.3%)	34 (13.7%)	0.460 (0.307–0.688)	<0.001
Atrial fibrillation					
–	584	479 (82.0%)	105 (18.0%)		
+	294	204 (69.4%)	90 (30.6%)	2.013 (1.453–2.788)	<0.001
Renal dysfunction					
–	748	616 (82.4%)	132 (17.6%)		
+	130	67 (51.5%)	63 (48.5%)	4.388 (2.965–6.494)	<0.001
Mitral valve insufficiency					
–	47	39 (83.0%)	8 (17.0%)		
Mild	463	373 (80.6%)	90 (19.4%)	1.176 (0.531–2.604)	0.689
Moderate	290	222 (76.6%)	68 (23.4%)	1.493 (0.666–3.349)	0.331
Severe	78	49 (62.8%)	29 (37.2%)	2.885 (1.187–7.016)	0.019
Left ventricular mass index (g/m <sup>2</sup> )	878	123.89 ± 34.32	137.19 ± 45.25	1.009 (1.005–1.013)	<0.001

Abbreviations: OR, odds ratio; CI, confidence interval; PCI, percutaneous coronary intervention; NT-proBNP, N-terminal pro-B-type natriuretic peptide.

estimation of latent trait models generally uses maximum likelihood estimators [15]. The expectation–maximization algorithm and quasi-Newton algorithm are commonly used in the iterative process. In this study, the mixed algorithm was used to estimate parameters. Therefore, the expectation–maximization algorithm was used to iterate at the beginning, and then quasi-Newton algorithm was used to iterate until convergence. Evaluation methods of a model include the likelihood ratio test, Pearson's test, the Akaike information criterion index, and the Bayesian information criterion index. Smaller values of the Akaike information criterion and Bayesian information criterion represent a better fit of the model [16–18]. In this study, the Akaike information criterion, Bayesian information criterion, and likelihood ratio test were used to compare model fitting. Additionally, the two-way residual was used to judge

whether the model fit well. After the optimal model was determined, the observation value was substituted into the model to obtain the prediction value of the individual latent trait score. By comparing the latent trait scores of the death group and the survival group, we were able to verify whether the selected variables could divide patients with HFrEF into subpopulations (test discrimination of the selected variables).

Finally, the selected variables were input into a nomogram model. After a simple sum of the scores from the nomogram model, the ABC-PI was created [19] (univariate analysis:  $\alpha = 0.05$ , multivariate analysis:  $\alpha_{\text{entry}} = 0.05$ , and  $\alpha_{\text{elimination}} = 0.10$ ). All statistical analyses were performed using SPSS 26.0 (<https://www.ibm.com/analytics/spss-statistics-software>) and R 3.6.1 (<https://www.r-project.org/>) software.



**Figure 1:** ROC curve. Abbreviations: ROC, receiver operating characteristic; LVMI, left ventricular mass index; NT-proBNP, N-terminal pro-B-type natriuretic peptide; NYHA, New York Heart Association.

**Ethics approval and informed consent:** The study was reviewed and approved by the Medical Ethics Committee of Shanxi Medical University, China (No. 2013099) and written informed consent was obtained from all participants.

### 3 Results

In the development group, diseases of these patients included coronary heart disease (388, 44.2%), old myocardial infarction (163, 18.6%), unstable angina pectoris (196, 22.3%), arrhythmia (83, 9.5%), and others (48, 5.5%). There were 195 (22.2%) deaths during the 3-year follow-up period. In the validation group, diseases of these patients included coronary heart disease (253, 38.9%), old myocardial infarction (147, 22.6%), unstable angina pectoris (113, 17.4%), arrhythmia (98, 15.1%), and others (40, 6.1%). There were 108 (16.6%) deaths during the 3-year follow-up period. The 222 baseline variables were tested one by one, and the standard basic variables and all variables that showed significant differences ( $p < 0.05$ ) are shown in Table 1.

Variables that were significantly different ( $p < 0.05$ ) in Table 1 were included in multivariate logistic regression analysis. The overall predictive capacity of the resulting model was good (81.1%), and the predictive capacity was greater in the survival group (93.7%) than in the death group (36.9%). To improve the predictive capacity and clinical application value of the model, an ROC curve was used to identify the

most sensitive cutoff value based on the results mentioned above (area under the curve ranged from 0.605 to 0.733; Figure 1). Additionally, continuous variables were converted into categorical variables. After conversion, the predictive capacity of the model was improved (81.9%). The effective prediction variables and cutoff value are shown in Table 2.

In the log-rank univariable test, all variables that showed significance in the logistic regression model passed the test. Multivariate Cox regression analysis was then performed. We found that not having had percutaneous coronary intervention (PCI), NT-proBNP levels  $>2,122$  ng/L, a left ventricular mass index  $>137.9$  g/m<sup>2</sup>, renal insufficiency, age  $>76.5$  years, NYHA grade IV, and atrial fibrillation were risk factors for death in patients with HFrEF. The model's prediction accuracy was high (C-index =  $0.801 \pm 0.013$ ; Table 3).

LTA was used to verify whether the selected variables could effectively distinguish the latent trait of HFrEF in patients and data were used from the validation group. Figure 2 shows the parameter estimation results of LTA, and Table 4 shows comparison of the two LTA models. The Ltm was better than the Rasch model. The Ltm was used to fit the latent trait scores and the latent trait score results were converted to a 10-point system. The scores of the survival and death groups were compared by the  $t$ -test. In the death group, the mean latent trait score was  $5.847 \pm 2.055$  and that in the survival group was  $3.086 \pm 2.228$  ( $t = 16.233$ ,  $p < 0.001$ ). This finding indicated that the selected seven variables with great discrimination could effectively distinguish the latent trait of HFrEF in patients.

**Table 2:** Multivariate logistic regression results after conversion of variables

Variable	Cutoff value	$\beta$	SE	Wald $\chi^2$	OR (95% CI)	P
Renal dysfunction		1.628	0.235	47.827	5.096 (3.212–8.085)	<0.001
PCI		−0.664	0.239	7.718	0.515 (0.322–0.822)	0.005
Atrial fibrillation		0.569	0.195	8.557	1.767 (1.207–2.588)	0.003
Log-NTproBNP (ng/L)	3.3268 (2122)	1.335	0.202	43.567	3.801 (2.557–5.650)	<0.001
NYHA grade	IV	0.664	0.202	10.767	1.943 (1.307–2.890)	0.001
Age	77	0.943	0.196	23.142	2.567 (1.748–3.769)	<0.001
LVMl	138	0.916	0.194	22.279	2.500 (1.709–3.658)	<0.001
Constant		−3.289	0.233	199.251	5.096 (3.212–8.085)	<0.001

Abbreviations: PCI, percutaneous coronary intervention; NT-proBNP, N-terminal pro-B-type natriuretic peptide; NYHA, New York Heart Association; LVMl, left ventricular mass index.

Nomogram scores were examined using R software and the highest total score was 415. For ease of application, nomogram scores were converted to percentiles. The 3-year survival rate was >42% in patients with total scores >30 points and it was >82% in those with total scores >60 points (Figure 3). Table 5 shows accumulation of the total score of 100 points. NT-proBNP levels had the highest predictive value for death with a score of 24. NYHA grade had a low predictive value, with 8 points (for better application, cutoff values of the variables were rounded down to the nearest integer).

To further validate the ABC-PI, 651 patients in the validation group were tested. Higher scores were associated with lower 1-year, 2-year, and 3-year death rates (Figure 4). The 3-year mortality rate was 34.7% in patients with scores of  $\leq 45$ , and it was only 2.6% in patients with scores of 68–100. The PI showed good performance.

## 4 Discussion

The prognosis of HFREF differs substantially depending on subpopulations. Therefore, accurate assessment of

this prognosis is important. In this study, we developed and validated the ABC-PI. The ABC-PI is a biomarker-based PI. The seven risk factors with the greatest capacity to predict risk of death in patients with HFREF were NT-proBNP, renal dysfunction, age, left ventricular mass index, PCI, atrial fibrillation, and NYHA grade. After verification by LTA, the selected index was considered to be valid. The score was formed by a nomogram and then converted to a percentage system, and the ABC-PI was finally obtained. After external validation, the PI was found to be effective and practical.

The ABC indicator had a high accuracy because of inclusion of age, a biomarker, and clinical history. A good prognostic indicator can be used in many aspects, especially to evaluate the effectiveness of medicine and precise treatment. Hijazi *et al.* used the ABC score to determine the effectiveness of anticoagulants in patients with atrial fibrillation [20]. The ABC-PI for patients with HFREF has not been determined yet. The ABC-PI used in our study could be useful for similar medical research.

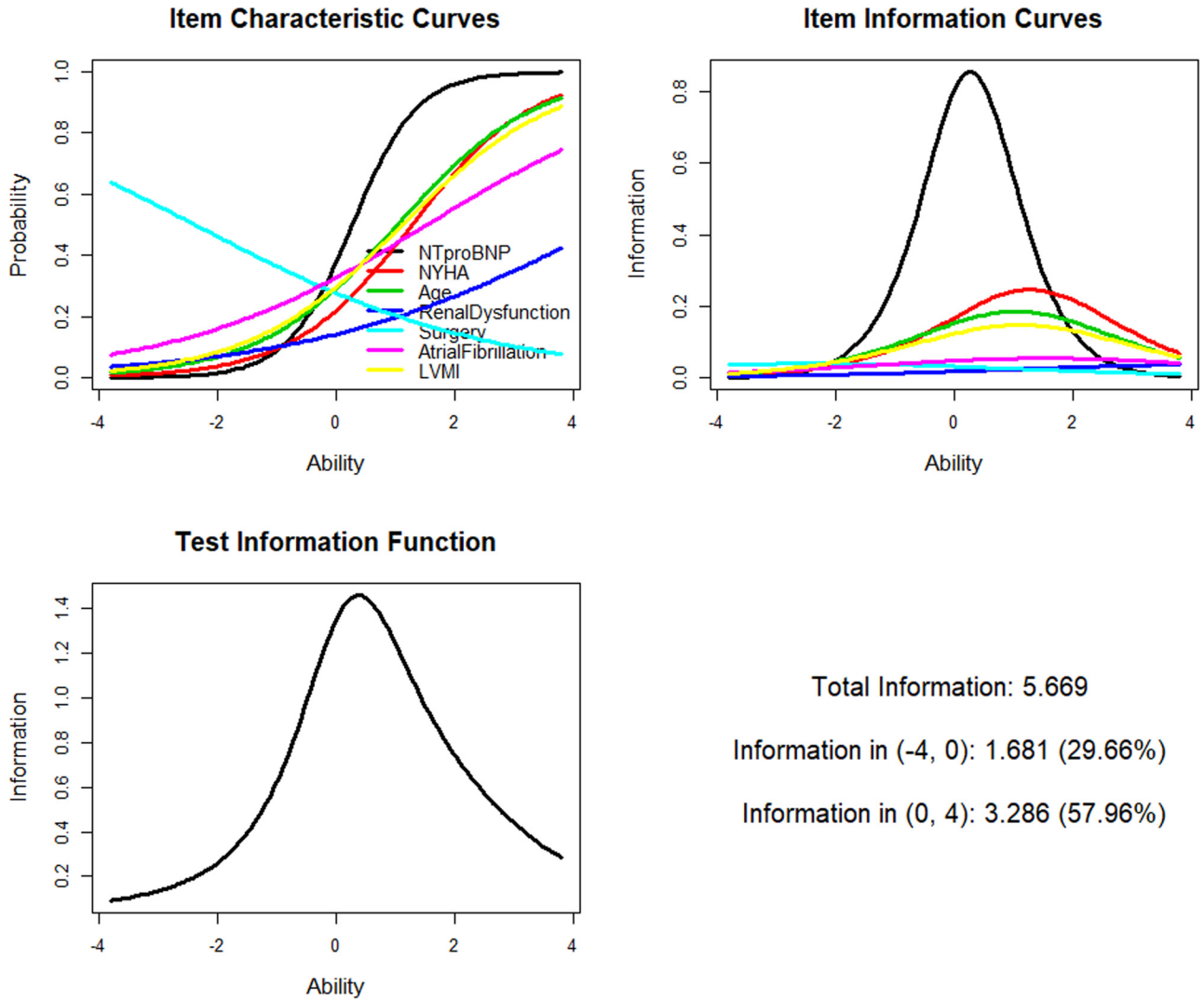
**Table 3:** Multivariate Cox regression

Variable	$\beta$	SE	Wald $\chi^2$	HR (95% CI)	P
PCI	−0.526	0.190	7.628	0.591 (0.407–0.858)	0.006
Log-NTproBNP (ng/L)	1.120	0.173	41.904	3.066 (2.184–4.304)	<0.001
LVMl	0.737	0.148	24.668	2.090 (1.563–2.797)	<0.001
Renal dysfunction	1.022	0.155	43.302	2.780 (2.050–3.769)	<0.001
Age	0.582	0.150	14.976	1.789 (1.332–2.401)	<0.001
NYHA grade	0.439	0.150	8.527	1.551 (1.155–2.082)	0.003
Atrial fibrillation	0.426	0.147	8.424	1.532 (1.149–2.043)	0.004

Notes: C-index:  $0.801 \pm 0.013$ .

Abbreviations: PCI, percutaneous coronary intervention; LVMl, left ventricular mass index; NT-proBNP, N-terminal pro-B-type natriuretic peptide; NYHA, New York Heart Association.





**Figure 2:** LTA curve. Abbreviations: LTA, latent trait analysis; LVMI, left ventricular mass index; NT-proBNP, N-terminal pro-B-type natriuretic peptide; NYHA, New York Heart Association.

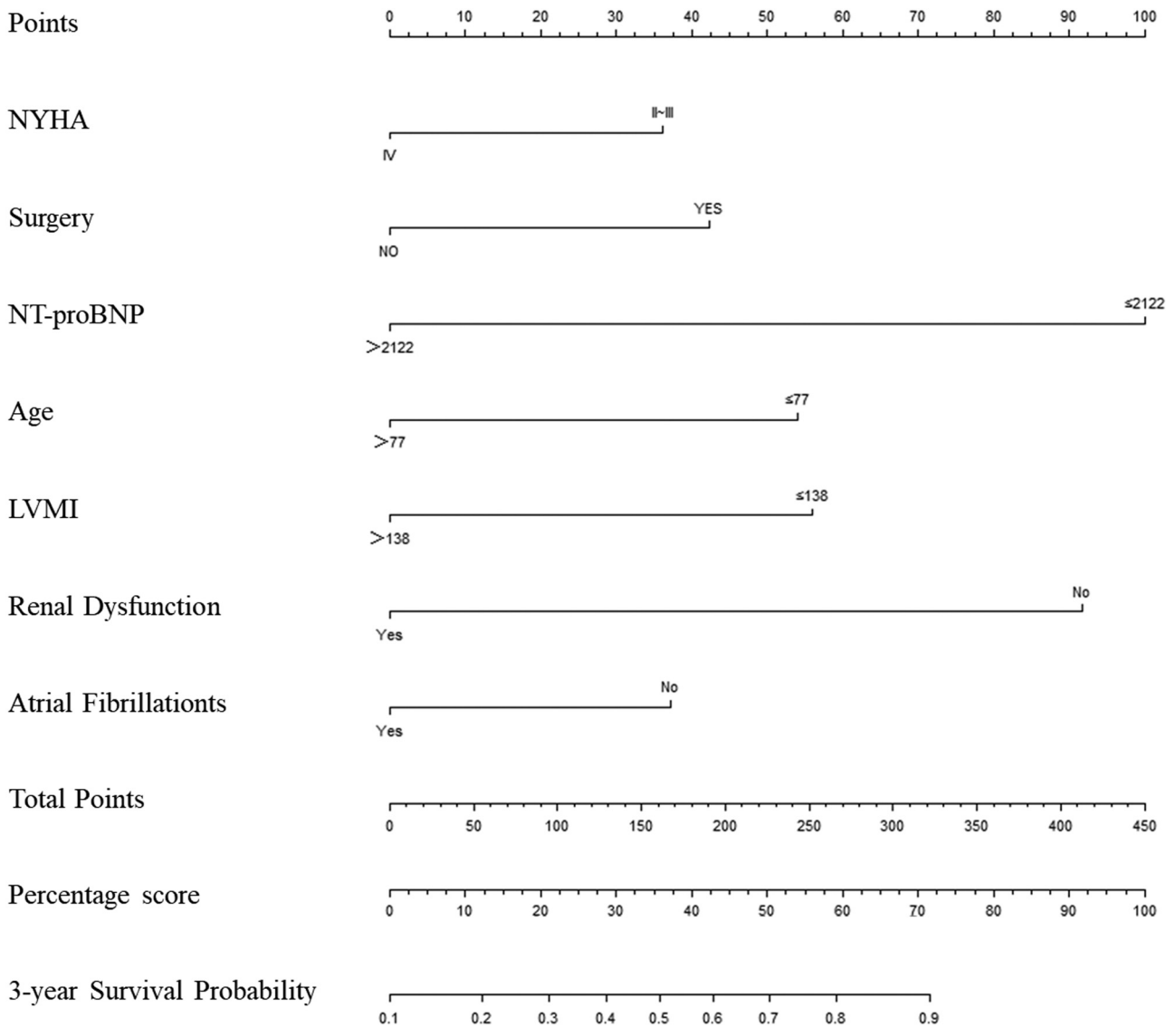
**Table 4:** Comparison of Ltm and Rasch model fitting indices in LTA

Model	AIC	BIC	Log.Lik	LRT	df	p value
Rasch	7319.76	7353.20	-3652.88			
Ltm	7253.80	7292.02	-3618.90	67.96	1	<0.001

Abbreviations: AIC, akaike information criterion; BIC, Bayesian information criterion, LRT, likelihood ratio test; df, degree of freedom.

The most informative predictor in the current study was NT-proBNP, which is a common biomarker of heart failure. Other biomarkers that have been identified include ST2 and galectin-3, but they are not routinely measured variables, and there is no evidence that their predictive capacity is better than that of NT-proBNP [21,22]. Therefore, NT-proBNP was used in the final PI in our study. Renal

dysfunction is one of the most common comorbidities in patients with HFrEF. In a Swedish study, in patients with a reduced ejection fraction, those with renal dysfunction had a 1-year mortality rate of 23%, whereas those without renal dysfunction had a mortality rate of 8% [23]. In the previous study, after adjusting for other factors, the hazard ratio was 1.51 and the 95% confidence interval was 1.40–1.63. The hazard ratio of renal dysfunction in the present study was 2.78 (95% confidence interval: 2.05–3.77), which is similar to those of previous studies [24,25]. Age has a well-known important effect on the prognosis of HFrEF. A cutoff of 60 years was used when the variable of age was used to predict heart failure in some previous studies [25]. However, since the subsequent development of new medicines and surgical techniques, heart failure currently tends to progress more slowly. The life span of patients with heart



**Figure 3:** Nomogram. Abbreviations: NYHA, New York Heart Association; NT-proBNP, N-terminal pro-B-type natriuretic peptide; LVMI, left ventricular mass index.

failure has become prolonged, and accordingly, recent studies have determined that an age of 60 years is not an ideal cutoff. The best cutoff age may be between 67 and 80 years [26,27]. In the current study, with regard to the ROC curve, the optimal cutoff age was 77 years. This age may constitute a better cutoff for use in future studies. The left ventricular mass index is a combination of left ventricular mass and baseline body surface area. This index is associated with a variety of heart diseases and prognoses, including coronary heart disease, structural heart disease, and heart failure [28–30]. The left ventricular mass index also showed good predictive value for the prognosis of patients with HFrEF in the present study.

Traditionally, a PI that is generated on the basis of Cox regression is highly dependent on the model and requires

re-input in the model. Based on the traditional Cox regression, this study used LTA to verify discrimination of the index. We then used a nomogram to form the ABC-PI, which did not depend on the original model, and it improved the practicability and applicability of the PI. LTA is commonly used in psychological tests and educational tests. LTA can distinguish different subpopulations of people [31]. Our study applied LTA to different latent traits with HFrEF, and it successfully validated discrimination of the selected variables. In the future, LTA could be used in similar diseases.

The current study has some limitations. The main limitation is that because the study population was limited to China, the results cannot necessarily be directly extrapolated to other populations. In future studies, we

**Table 5:** Nomogram scores (centesimal system)

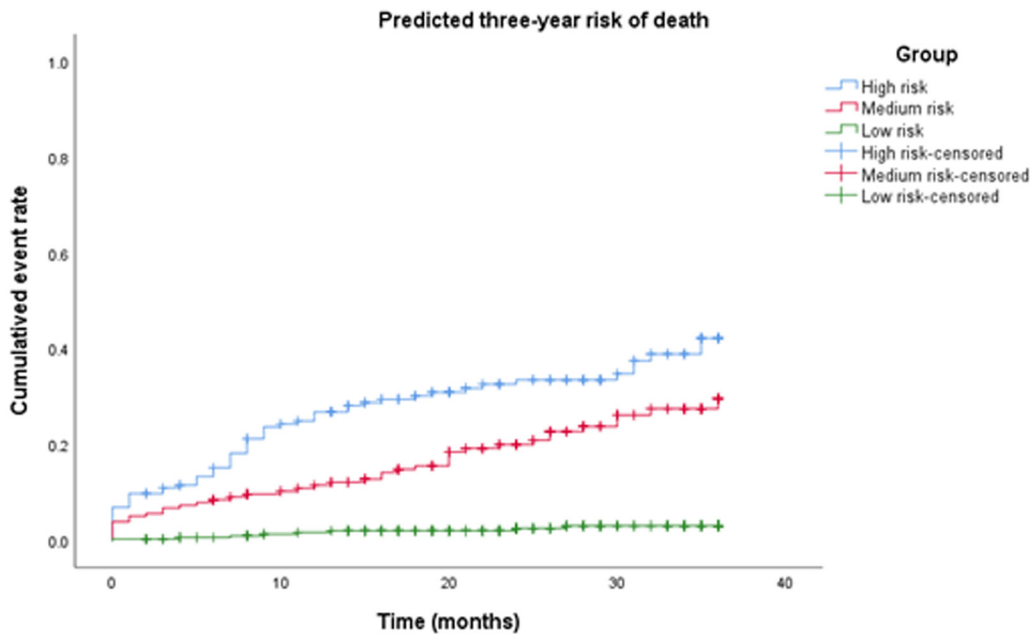
Variable	Category	Score
NT-proBNP (ng/L)	≤2,100	24
	>2,100	0
Renal insufficiency	–	22
	+	0
Age (years)	≤77	13
	>77	0
LVMI (g/m <sup>2</sup> )	≤140	13
	>140	0
PCI	–	0
	+	10
Atrial fibrillation	–	10
	+	0
NYHA grade	II/III	8
	IV	0
	Total	100

Abbreviations: LVMI, left ventricular mass index; NT-proBNP, N-terminal pro-B-type natriuretic peptide; PCI, percutaneous coronary intervention; NYHA, New York Heart Association.

will attempt to incorporate heart failure research centers in other countries to further improve the accuracy of the ABC-PI.

## 5 Conclusion

Overall, the PI developed in the present study shows good performance. The variables that constitute the ABC-PI include age, biomarkers, and clinical history. Therefore, the scope of this index is comprehensive. After external validation, the ABC-PI performed well for predicting 1-year, 2-year, and 3-year mortality. The 1-year, 2-year, and 3-year mortality rates in the high-risk group were 26.0%, 31.2%, and 34.7% compared with 1.7%, 2.3%, and 2.6% in the low-risk group, respectively. In the future, the ABC-PI could be used for evaluating the effect of intervention and individual risk prediction.



Group (scores)	One-year mortality	Two-year mortality	Three-year mortality
High risk (0-45)	26.0%	31.2%	34.7%
Medium risk (46-67)	11.4%	18.2%	22.7%
Low risk (68-100)	1.7%	2.3%	2.6%

**Figure 4:** Cumulative risk of death predicted by the ABC-PI for different groups (validation group, *n* = 651). Abbreviations: ABC-PI, age-biomarker-clinical history prognostic index.



## Abbreviations

ABC-PI	age-biomarker-clinical history prognostic index
eCRF	electronic Case Report Form
HFrEF	reduced left ventricular ejection fraction
LTA	latent trait analysis
NT-proBNP	N-terminal pro-B-type natriuretic peptide
NYHA	New York Heart Association
PCI	percutaneous coronary intervention
PI	prognostic index
ROC	receiver operating characteristic.

**Acknowledgments:** The authors thank all patients, doctors, nurses, and investigators who were involved in the study and the touch-up editor. The authors also thank Knapp, PhD, from Liwen Bianji, Edanz Group China ([www.liwenbianji.cn/](http://www.liwenbianji.cn/)) for editing the English text of a draft of this manuscript.

**Conflict of interest:** None.

**Data availability:** All relevant non-patient-identifying data are contained within the published manuscript.

**Funding:** This research was supported by a grant from the National Natural Science Foundation of China (grant number 81872714) and Graduate Science and Technology Innovation projects organized by the Department of Education of Shanxi Province (grant number 2019SY272).

**Authors' contributions:** All authors participated in the design of the study. HL participated in researcher coordination, data analysis, and drafting of the manuscript. JT and QZ participated in eCRF development. CY and HY were primarily responsible for data collection and analysis, and KW participated in data analysis. YBZ and QHH conceptualized the study, supervised data analysis, and oversaw the generation of the original manuscript. All authors have read and approved the final manuscript.

## References

- [1] Bloom MW, Greenberg B, Jaarsma T, Januzzi JL, Lam CSP, Maggioni AP, et al. Heart failure with reduced ejection fraction. *Nat Rev Dis Primers*. 2017;3:17058. doi: 10.1038/nrdp.2017.58.
- [2] Braunwald E. Heart failure. *JACC Heart Failure*. 2013;1:1–20. doi: 10.1038/clpt.2013.149.
- [3] Yancy CW, Jessup M, Bozkurt B, Masoudi FA, Wilkoff BL. American College of Cardiology Foundation, American Heart Association Task Force on Practice Guidelines. 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. *J Am Coll Cardiol*. 2013;62:e147–239. doi: 10.1016/j.jacc.2013.05.020.
- [4] Lovic D, Stojanov V, Jakovljević B, Krotin M, Jurisic V, Djordjevic D, et al. Prevalence of arterial hypertension in Serbia: PAHIS study. *J Hypertens*. 2013;31(11):2151–7; discussion 2157. doi: 10.1097/HJH.0b013e328364c2a2.
- [5] Lip GYH, Lane DA. Assessing bleeding risk in atrial fibrillation with the HAS-BLED and ORBIT scores: clinical application requires focus on the reversible bleeding risk factors. *Eur Heart J*. 2015;36(46):3265–7. doi: 10.1093/eurheartj/ehv415.
- [6] Pencina MJ, D'Agostino RB, Larson MG, Massaro JM, Vasan RS. Predicting the 30-year risk of cardiovascular disease: the Framingham heart study. *Circulation*. 2009;119(24):3078–84. doi: 10.1161/CIRCULATIONAHA.108.816694.
- [7] Patoulias D, Stavropoulos K, Imprialos K, Athyros V, Grassos H, Doumas M, et al. Inflammatory markers in cardiovascular disease; lessons learned and future perspectives. *Curr Vasc Pharmacol*. 2020. doi: 10.2174/1570161118666200318104434. [preprint].
- [8] Chalikiopoulou C, Bizjan BJ, Leventopoulos G, Smaili K, Blagus T, et al. Multiomics analysis coupled with text mining identify novel biomarker candidates for recurrent cardiovascular events. *OMICS*. 2020;24(4):205–15. doi: 10.1089/omi.2019.0216.
- [9] Farnsworth CW, Bailey AL, Jaffe AS, Scott MG. Diagnostic concordance between NT-proBNP and BNP for suspected heart failure. *Clin Biochem*. 2018;59:50–5. doi: 10.1016/j.clinbiochem.2018.07.002.
- [10] C W. Nomenclature and Criteria for Diagnosis of Diseases of the Heart. Little, Brown. 1979. doi: 10.1001/jama.1940.02810200082036.
- [11] Sugumar H, Nanayakkara S, Prabhu S, Voskoboinik A, Kaye DM, Ling LH, et al. Pathophysiology of atrial fibrillation and heart failure: dangerous Interactions. *Cardiol Clin*. 2019;37:131–8. doi: 10.1016/j.ccl.2019.01.002.
- [12] Lazarsfeld PF, Henry NW. Latent structure analysis. *Am Sociol Rev*. 1968;34(2):293–94. doi: 10.2307/2092222.
- [13] Bock RD, Aitkin M. Marginal maximum likelihood estimation of item parameters: application of an EM algorithm. *Psychometrika*. 1981;46:443–59. doi: 10.1007/bf02293801.
- [14] Rasch G. Studies in mathematical psychology: I. Probabilistic models for some intelligence and attainment tests. 1960.
- [15] David J. Latent variable models and factor analysis: a unified approach, 3rd edition. *Int Stat Rev*. 2013;81(2):333–334. doi: 10.1198/tech.2001.s568.
- [16] Gollini I, Murphy TB. Mixture of latent trait analyzers for modelbased clustering of categorical data. *Stat Comput*. 2013;24(4):569–88. doi: 10.1007/s11222-013-9389-1.
- [17] Choi I Model selection for factor analysis: some new criteria and performance comparisons. Working Papers, 2013.
- [18] Hirose K, Kawano S, Konishi S, Ichikawa M. Bayesian information criterion and selection of the number of factors in factor analysis models. *J Data Sci*. 2011;9(2):243–59. doi: 10.6339/JDS.2011.09(2).927.
- [19] Zhang Y, Xie H, Zhang Z, Zhang P, Chen P, Wang X. The characteristics and nomogram for primary lung papillary adenocarcinoma. *Open Med*. 2020;15(1):92–102. doi: 10.1515/med-2020-0014.
- [20] Hijazi Z, Oldgren J, Lindbäck J, Alexander JH, Connolly SJ, Eikelboom JW, et al. A biomarker-based risk score to predict

- death in patients with atrial fibrillation: the ABC (age, biomarkers, clinical history) death risk score. *Eur Heart J*. 2017;39(6):477–85. doi: 10.1093/eurheartj/ehx584.
- [21] Weinberg EO, Shimpo M, Hurwitz S, Tominaga S, Rouleau JL, Lee RT. Identification of serum soluble ST2 receptor as a novel heart failure biomarker. *Circulation*. 2003;107:721–6. doi: 10.1161/01.cir.0000047274.66749.fe.
- [22] Besler C, Lang D, Urban D, Rommel KP, Von RM, Fengler K. Plasma and cardiac galectin-3 in patients with heart failure reflects both inflammation and fibrosis: implications for its use as a biomarker. *Circ Heart Fail*. 2017;10(3):e003804. doi: 10.1161/CIRCHEARTFAILURE.116.003804.
- [23] Löfman I, Szummer K, Dahlström U, Jernberg T, Lund LH. Associations with and prognostic impact of chronic kidney disease in heart failure with preserved, mid-range, and reduced ejection fraction. *Eur J Heart Fail*. 2017;19:1606–14. doi: 10.1002/ejhf.821.
- [24] Ter MJM, Damman K, Verhaar MC, Paulus WJ, Duncker DJ, Cheng C. Connecting heart failure with preserved ejection fraction and renal dysfunction: the role of endothelial dysfunction and inflammation. *Eur J Heart Fail*. 2016;18:588–98. doi: 10.1002/ejhf.497.
- [25] Warraich HJ, Kitzman DW, Whellan DJ, Duncan PW, Mentz RJ, Pastva AM. Physical function, frailty, cognition, depression, and quality of life in hospitalized adults  $\geq 60$  years with acute decompensated heart failure with preserved versus reduced ejection fraction. *Circ Heart Fail*. 2018;11(11):e005254. doi: 10.1161/CIRCHEARTFAILURE.118.005254.
- [26] Petrie MC, Jhund PS, She L, Adlbrecht C, Doenst T, Panza JA, et al. Ten-year outcomes after coronary artery bypass grafting according to age in patients with heart failure and left ventricular systolic dysfunction: an analysis of the extended follow-up of the STICH trial (surgical treatment for ischemic heart failure). *Circulation*. 2016;134:1314–24. doi: 10.1161/CIRCULATIONAHA.116.024800.
- [27] Berthelot E, Nouhaud C, Lafuente C, Assayag P, Hittinger L. Heart failure in patients over 80 years old. *Presse Med*. 2019;48:143–53. doi: 10.1016/j.lpm.2019.02.001.
- [28] Verma S, Mazer CD, Yan AT, Mason T, Garg V, Teoh H, et al. Effect of empagliflozin on left ventricular mass in patients with type 2 diabetes mellitus and coronary artery disease: the EMPA-HEART CardioLink-6 randomized clinical trial. *Circulation*. 2016;140:1693–702. doi: 10.2337/dc16-1312.
- [29] Chahal NS, Lim TK, Jain P, Chambers JC, Kooner JS, Senior R. New insights into the relationship of left ventricular geometry and left ventricular mass with cardiac function: a population study of hypertensive subjects. *Eur Heart J*. 2010;31:588–94. doi: 10.1093/eurheartj/ehp490.
- [30] Arenja N, Fritz T, Andre F, Riffel JH, Aus dem SF, Ochs M, et al. Myocardial contraction fraction derived from cardiovascular magnetic resonance cine images-reference values and performance in patients with heart failure and left ventricular hypertrophy. *Eur Heart J Cardiovasc Imaging*. 2017;18:1414–22. doi: 10.1093/ehjci/jew324.
- [31] Dayton CM. Latent class scaling analysis. (Quantitative Applications in the Social Sciences, Vol. 126. Newbury Park, California: Sage Publications; May 1999.