

1 **Derivation and Internal Validation of Prediction Models for**
2 **Pulmonary Hypertension Risk Assessment in a Cohort**
3 **Inhabiting Tibet, China**

4 **Authors**

5 Junhui Tang¹, Rui Yang², Hui Li¹, Xiaodong Wei¹, Zhen Yang¹, Wenbin Cai¹, Yao Jiang¹,
6 Ga Zhuo¹, Li Meng¹, Yali Xu³ *

7 Junhui Tang and Rui Yang contributed equally to this work.

8

9 **Affiliations**

- 10 1. Department of Ultrasound, the General Hospital of Tibet Military Area Command, Tibet,
11 China
- 12 2. Department of High Mountain Sickness, the General Hospital of Tibet Military Area
13 Command, Tibet, China
- 14 3. Department of Ultrasound, Xinqiao Hospital, Army Medical University, Chongqing, China

15

16 **Corresponding author ***

17 Yali Xu, MD, PhD

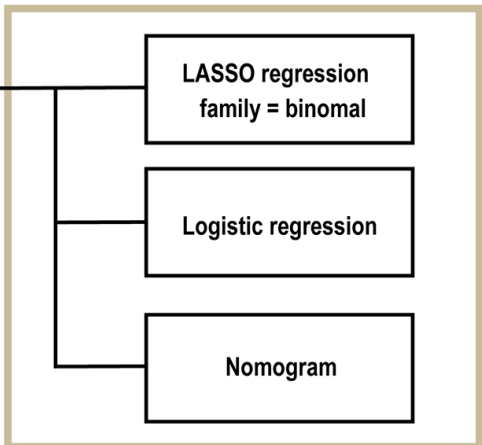
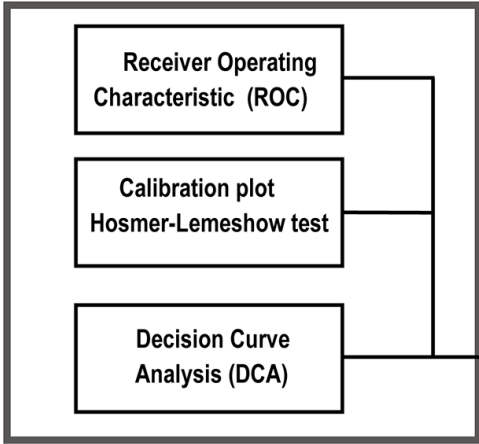
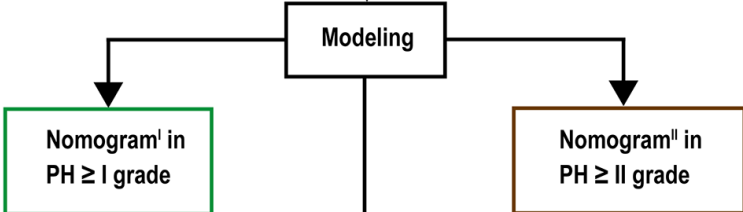
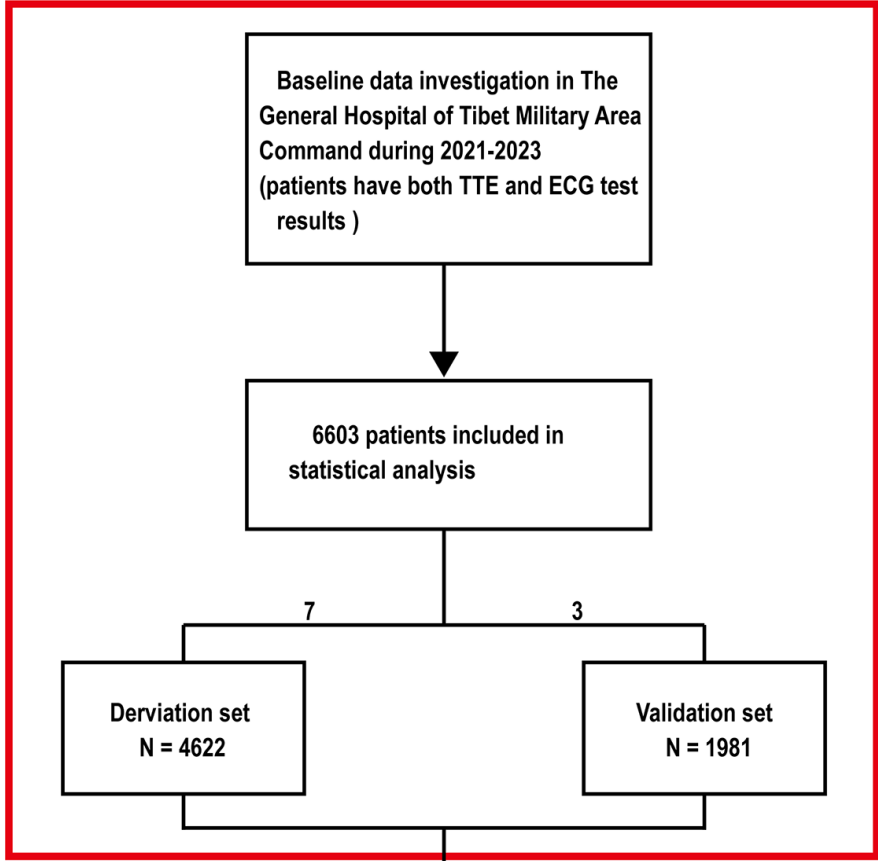
18 Department of Ultrasound, Xinqiao Hospital, Army Medical University, 83 Xinqiao Main Street,
19 Shapingba District, Chongqing, 400037, PR China.

20 Tel.: +86 18523384936

21 E-mail address: xuyali1976@163.com

22

23 This research was supported by the Talent Program of Army Medical University (No. 2019R038).

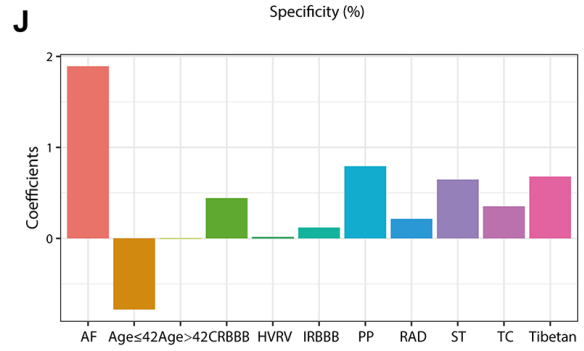
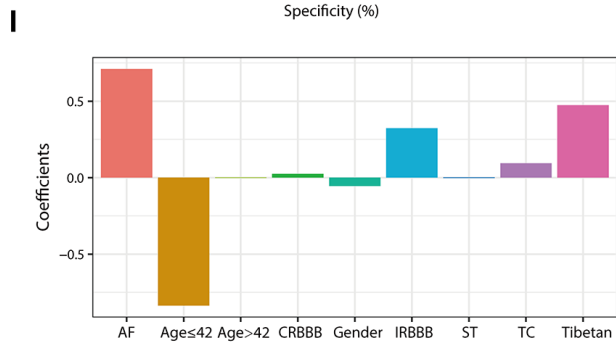
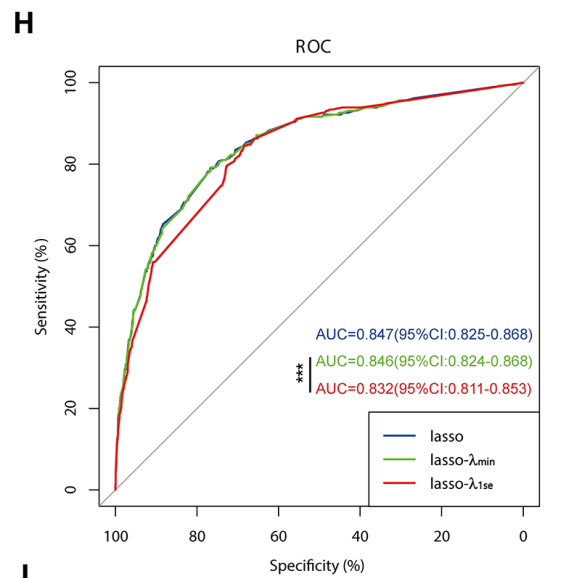
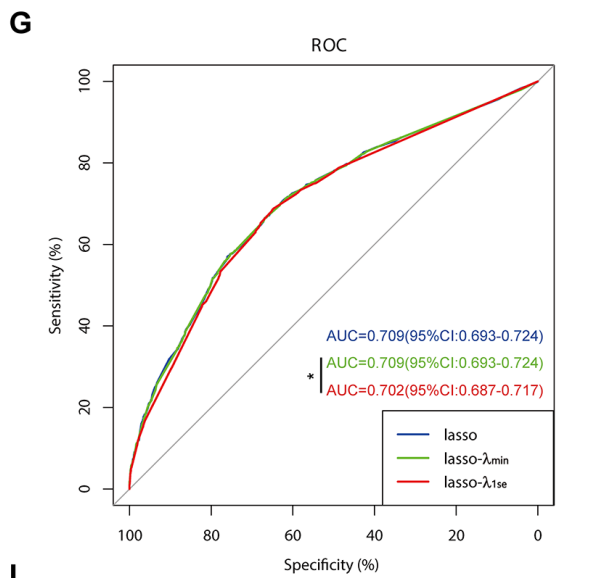
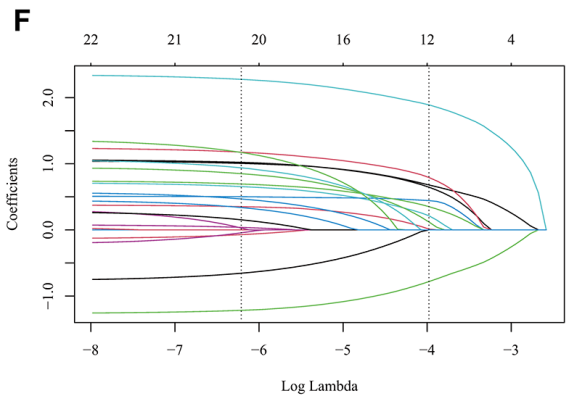
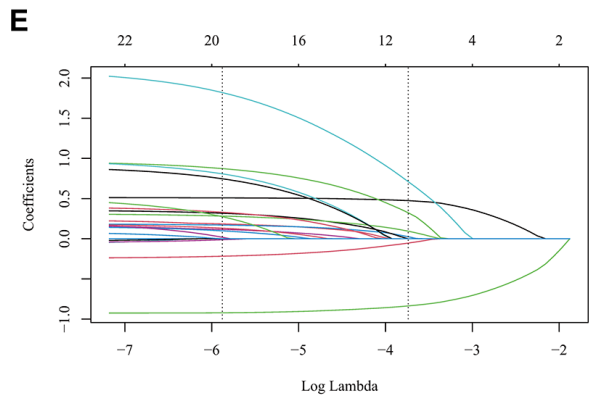
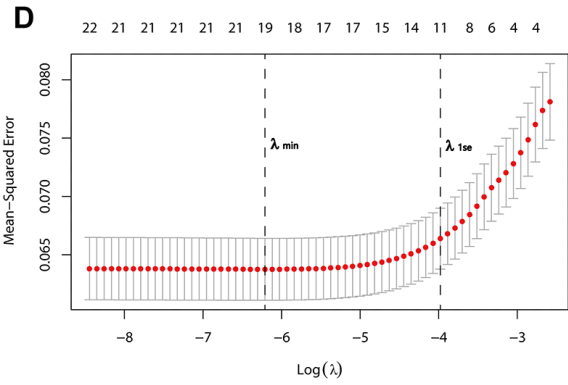
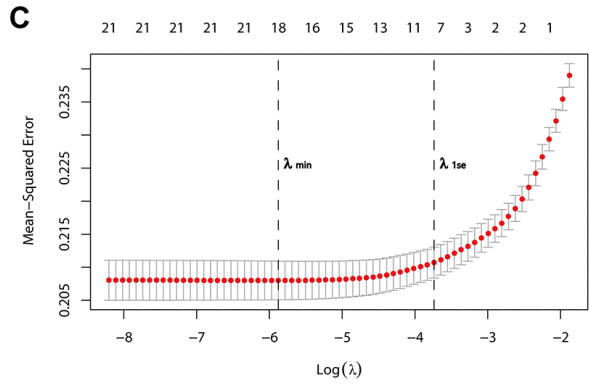
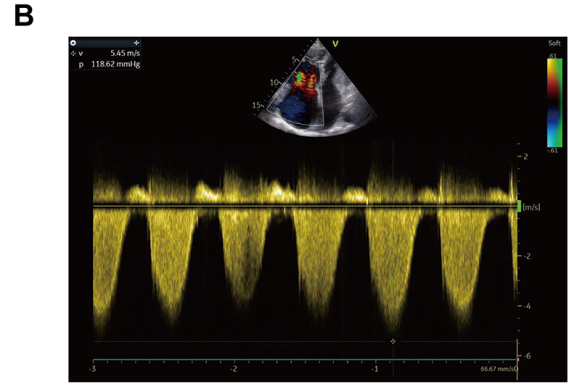
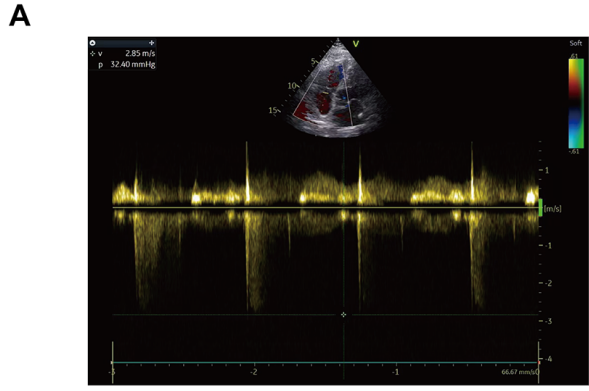


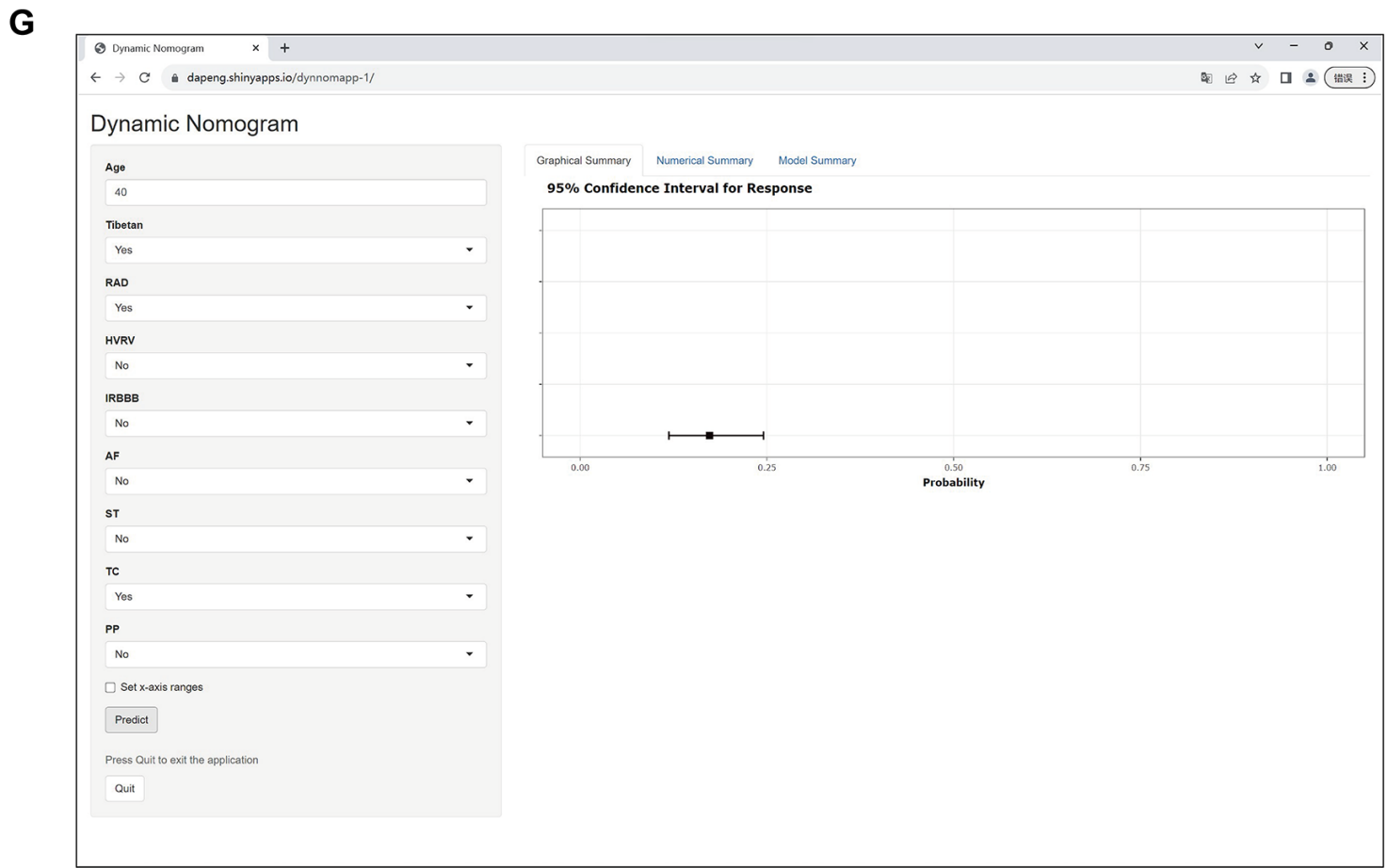
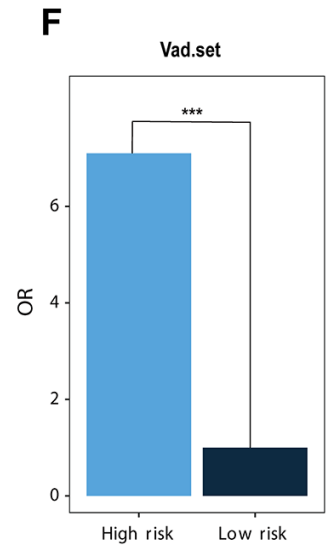
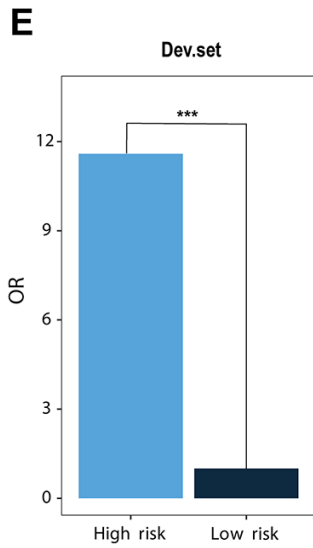
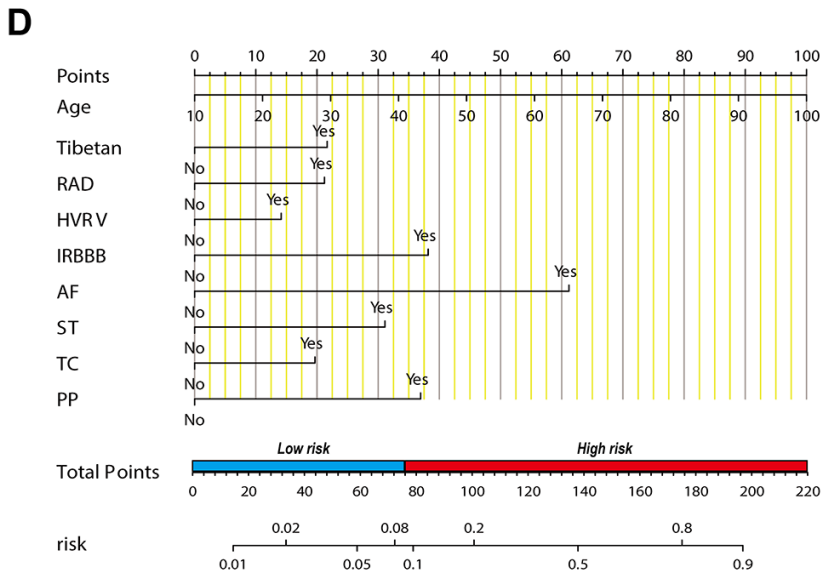
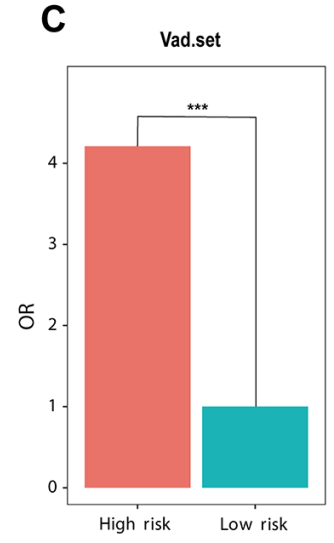
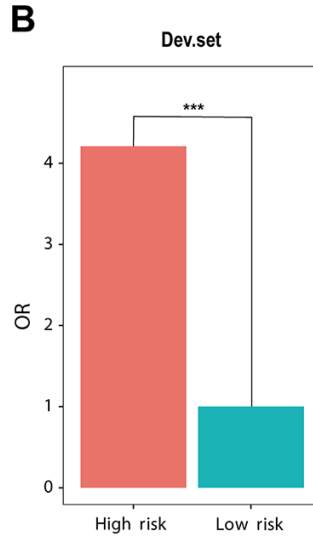
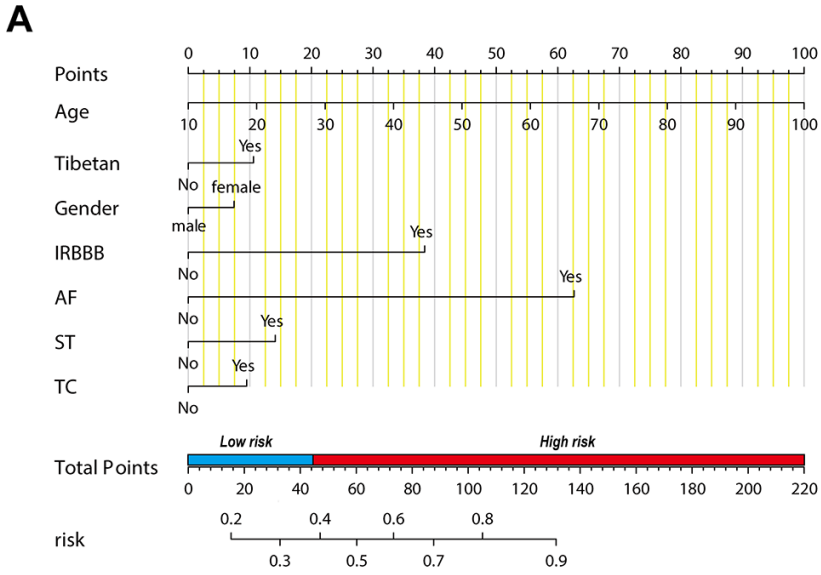
Introducing factors of Age, Tibetan ethnicity, Gender, IRBBB, AF, ST and TC, the risk nomogram^I is beneficial for prediction of PH ≥ I grade risk in population living in high altitude localities.

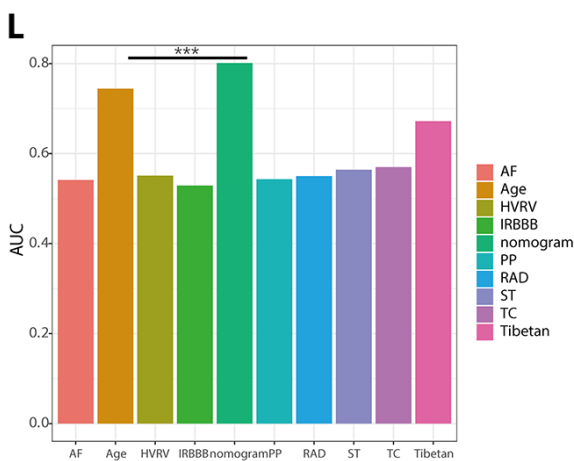
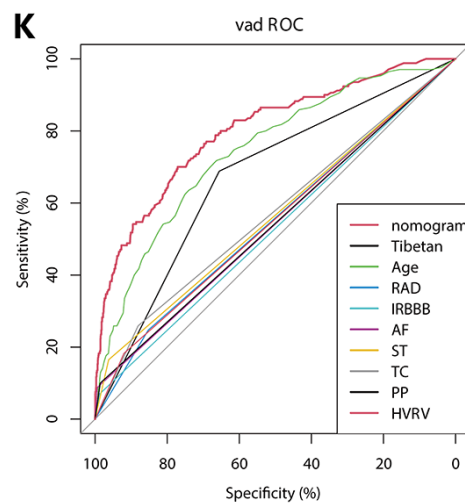
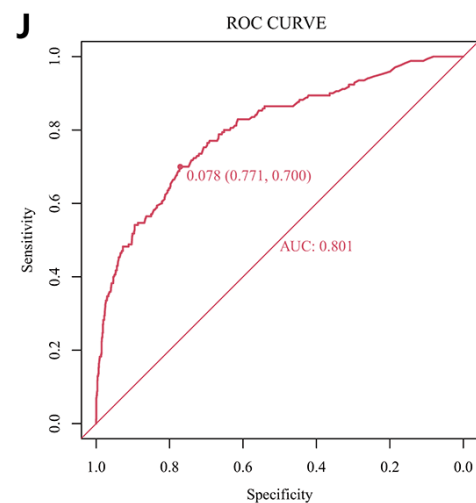
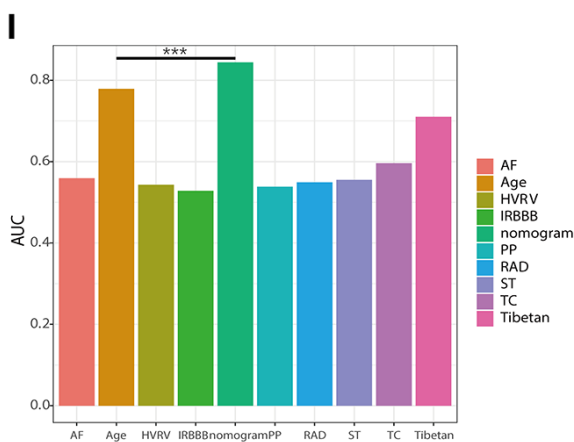
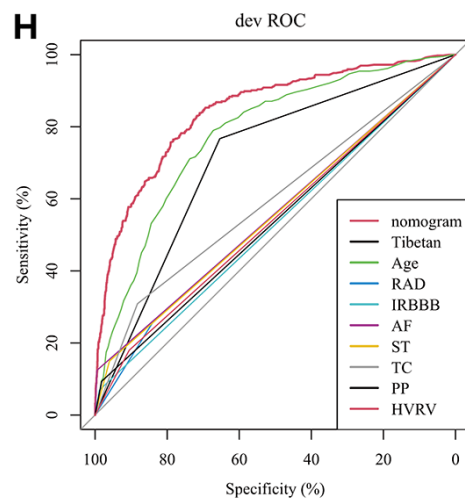
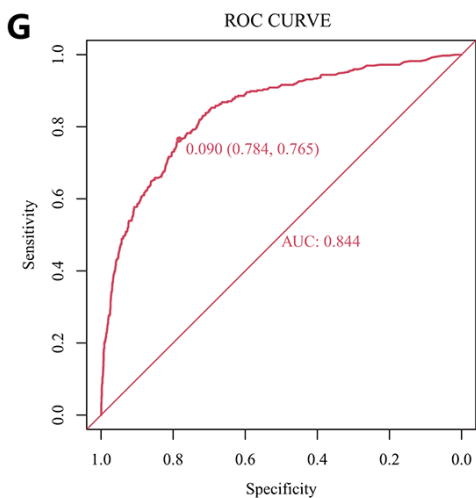
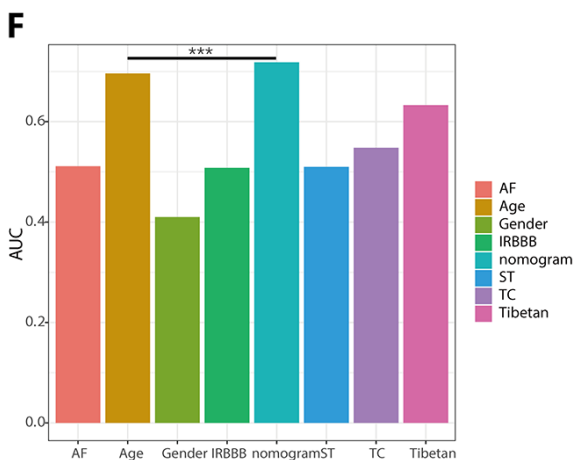
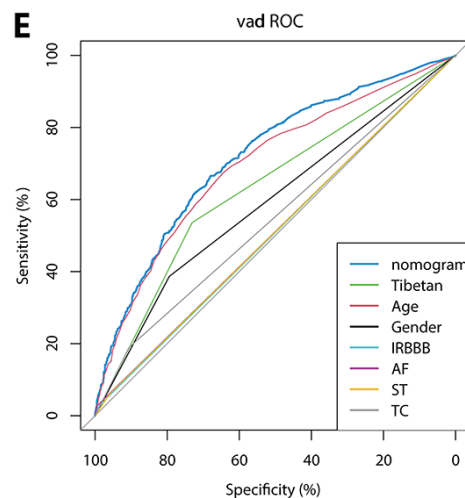
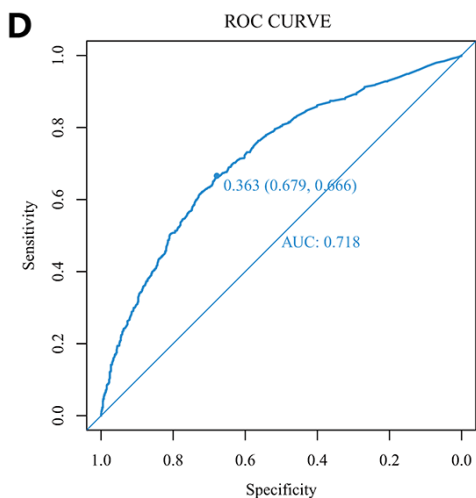
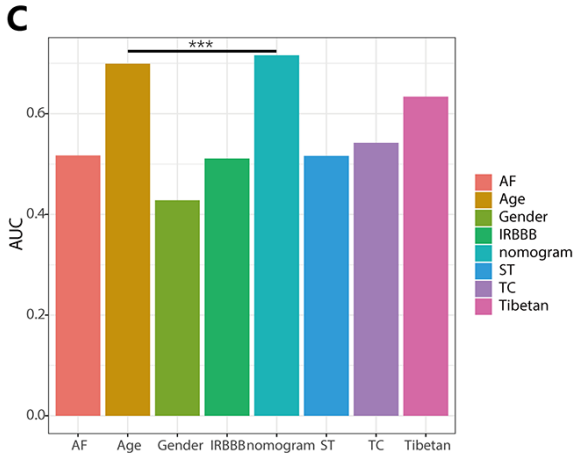
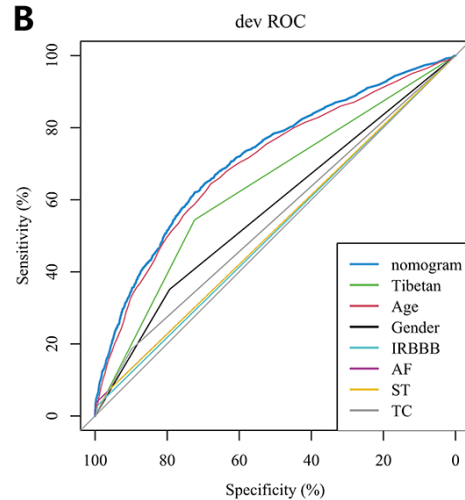
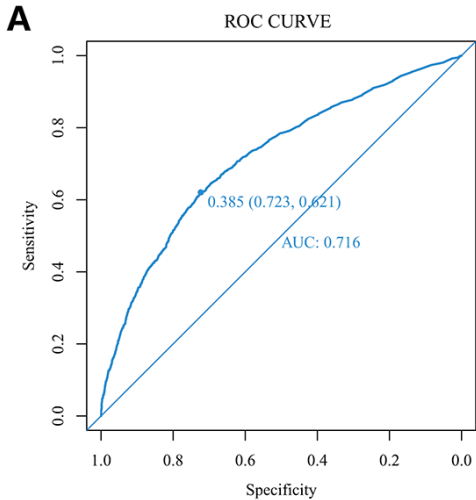
Introducing factors of Age, Tibetan ethnicity, RAD, HVRV, IRBBB, AF, ST, TC and PP, the risk nomogram^{II} is beneficial for prediction of PH ≥ II grade risk in population living in high altitude localities.

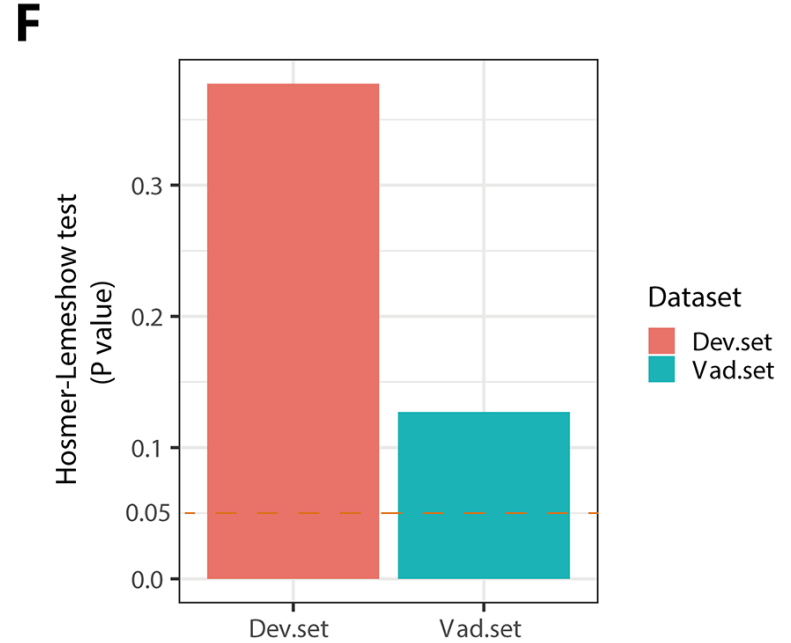
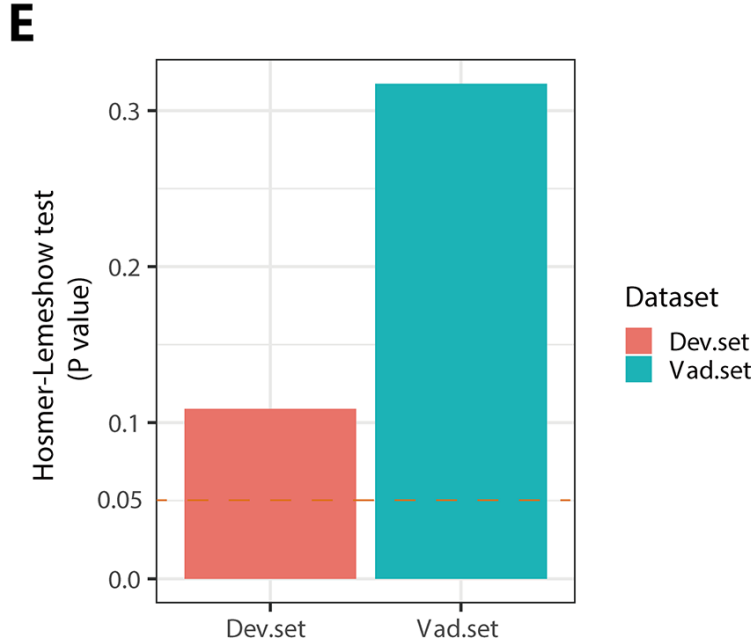
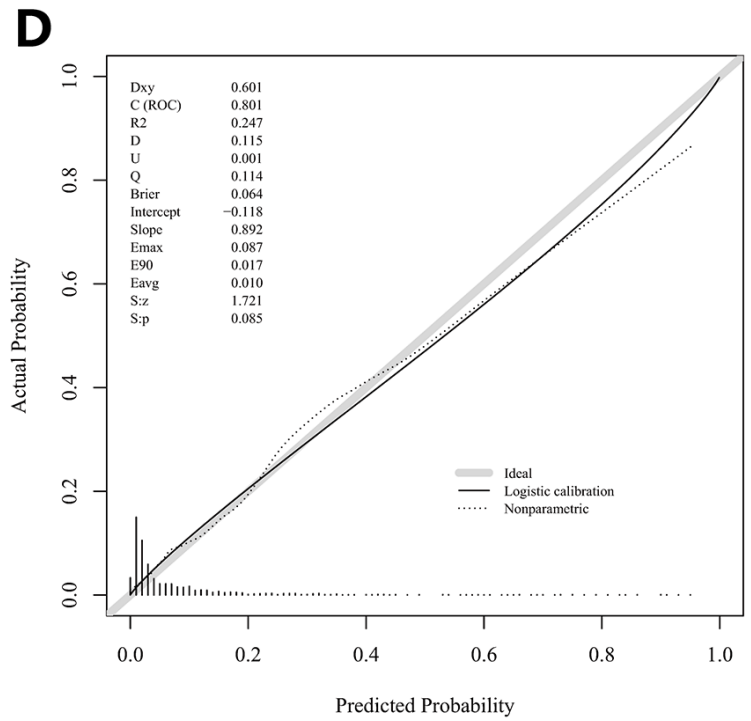
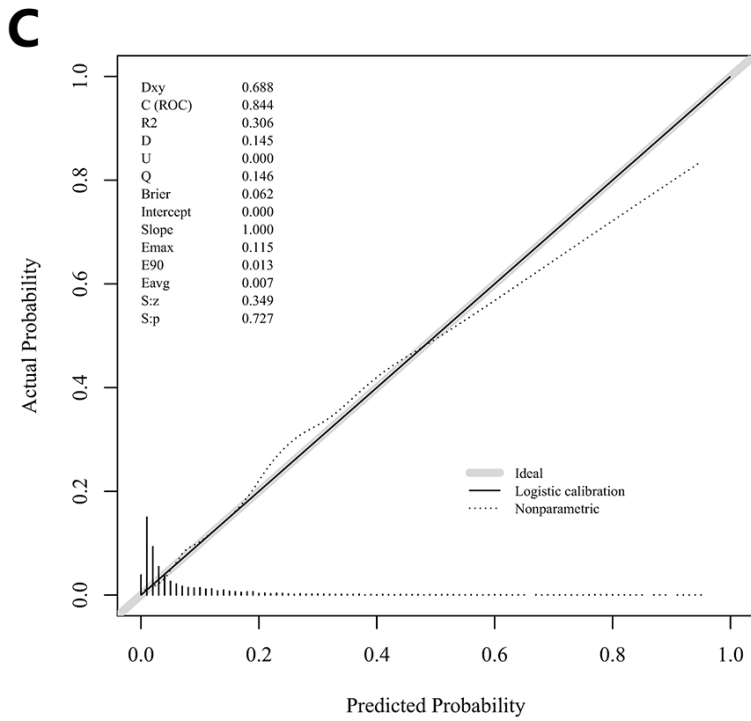
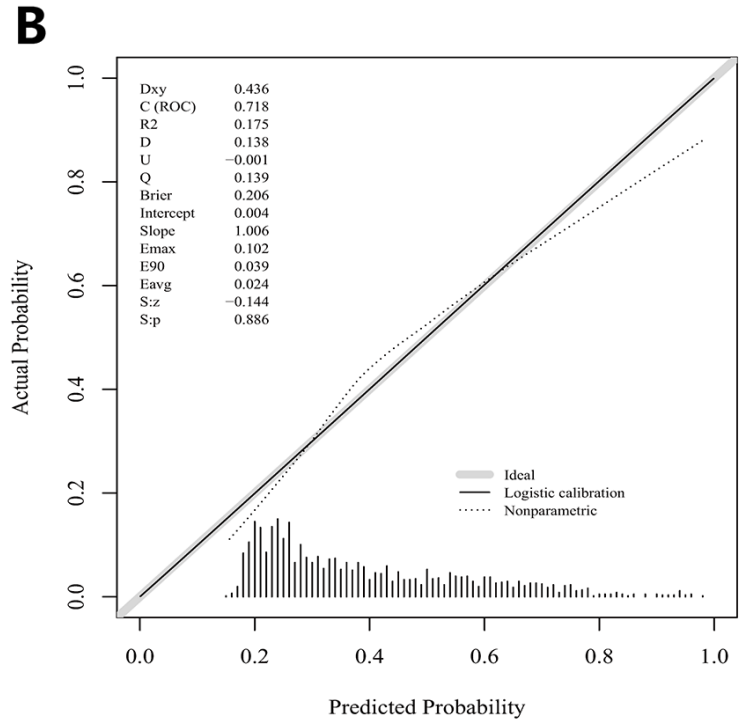
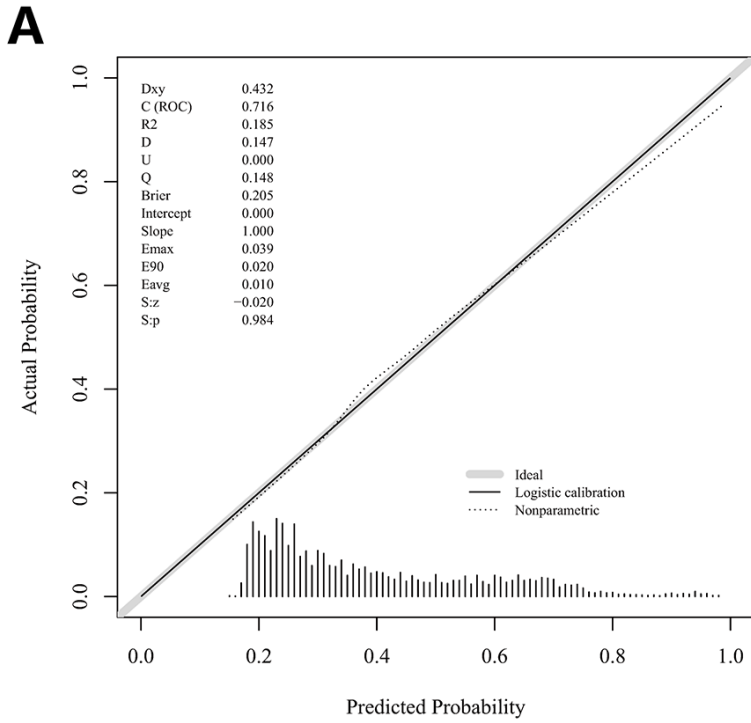
Results

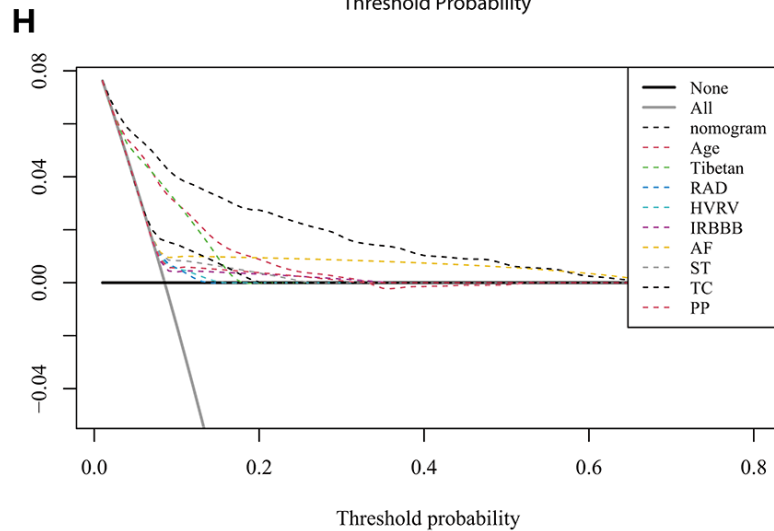
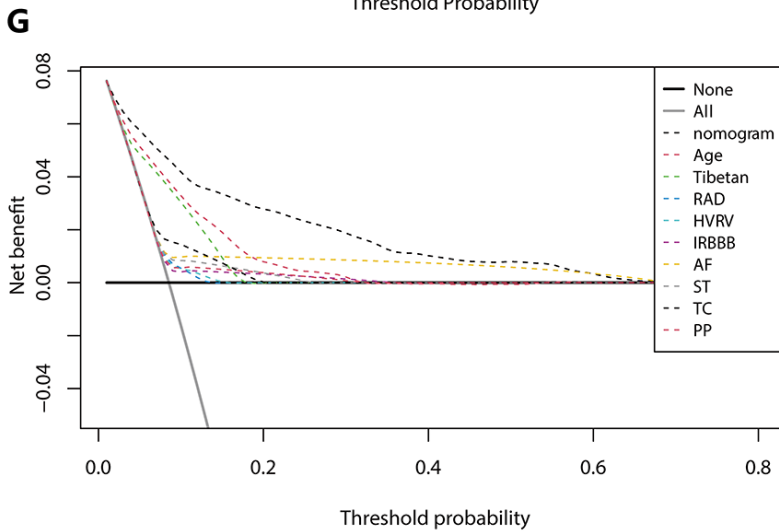
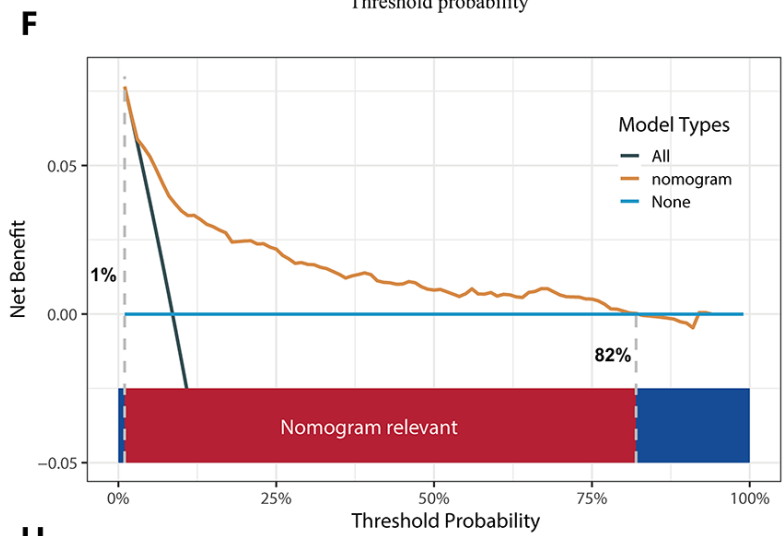
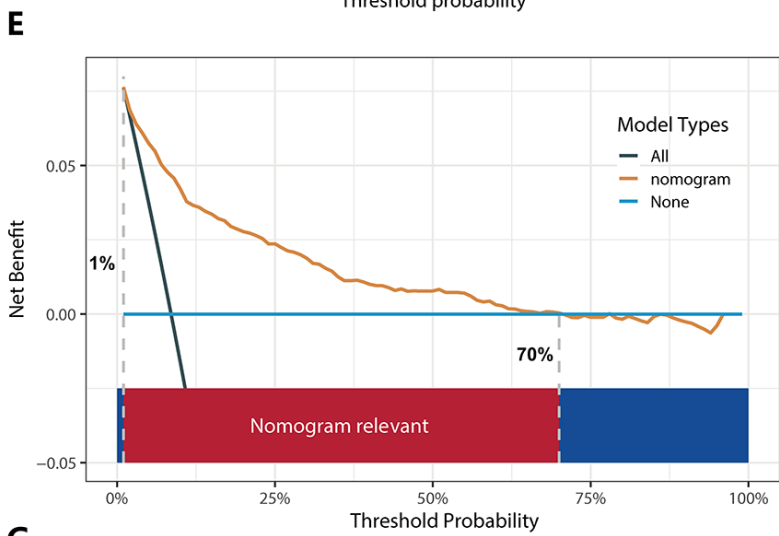
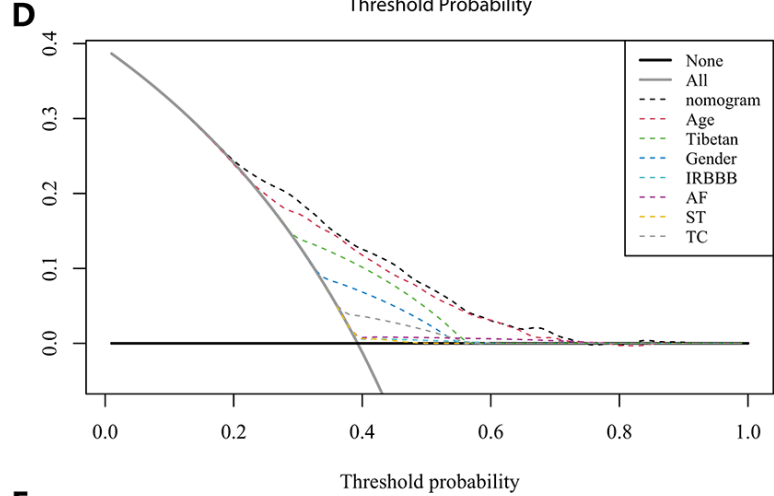
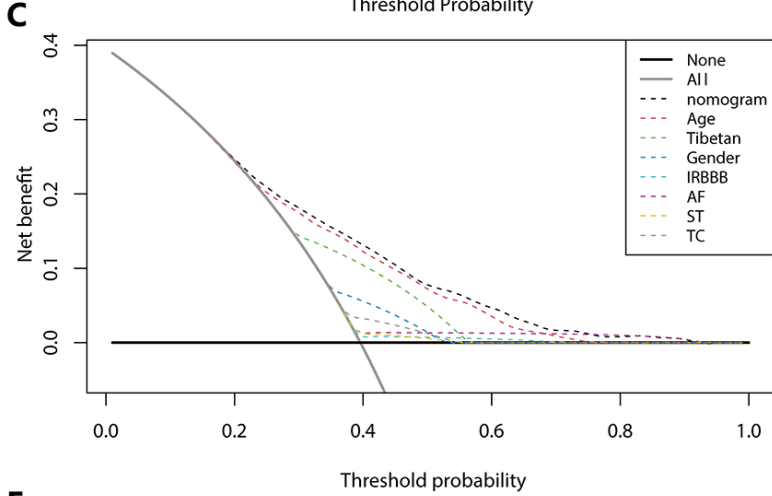
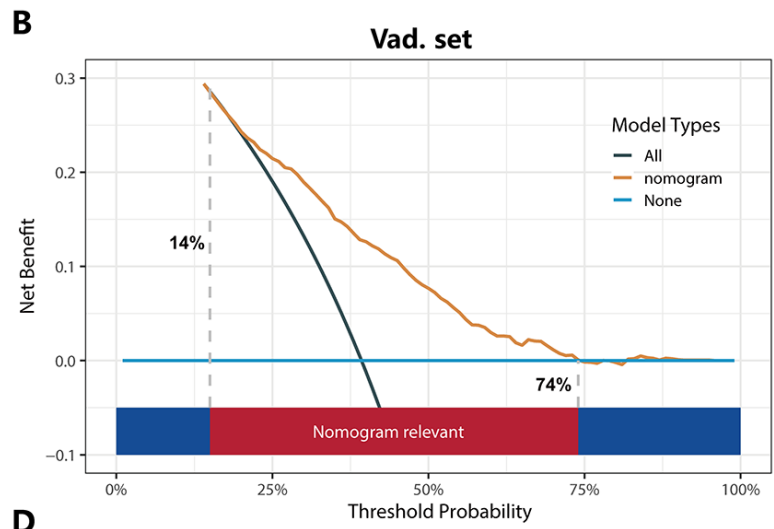
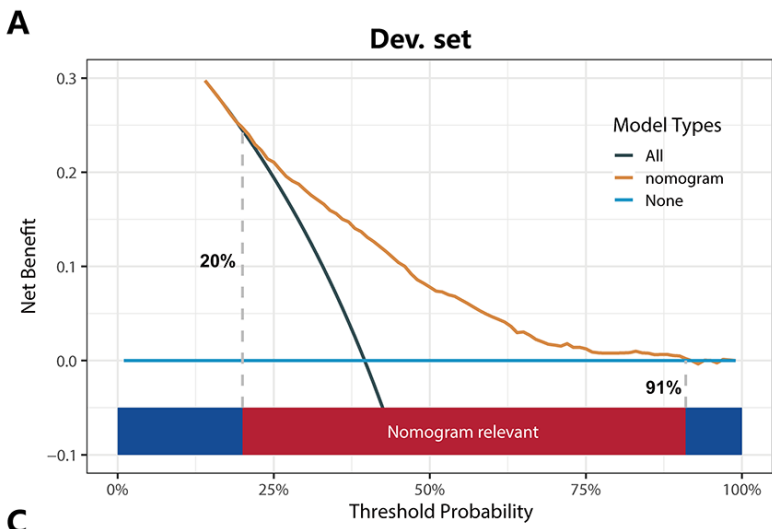
Conclusions
 The nomogram risk models of Nomogram^I and Nomogram^{II} are both well established and useful, and the model of Nomogram^{II} performs better.











1 **Derivation and Internal Validation of Prediction Models for**
2 **Pulmonary Hypertension Risk Assessment in a Cohort**
3 **Inhabiting Tibet, China**

4 **Authors**

5 Junhui Tang¹, Rui Yang², Hui Li¹, Xiaodong Wei¹, Zhen Yang¹, Wenbin Cai¹, Yao Jiang¹,
6 Ga Zhuo¹, Li Meng¹, Yali Xu^{3*}

7 Junhui Tang and Rui Yang contributed equally to this work.

8

9 **Affiliations**

- 10 1. Department of Ultrasound, the General Hospital of Tibet Military Area Command, Tibet,
11 China
- 12 2. Department of High Mountain Sickness, the General Hospital of Tibet Military Area
13 Command, Tibet, China
- 14 3. Department of Ultrasound, Xinqiao Hospital, Army Medical University, Chongqing, China

15

16 **Corresponding author ***

17 Yali Xu, MD, PhD

18 Department of Ultrasound, Xinqiao Hospital, Army Medical University, 83 Xinqiao Main Street,
19 Shapingba District, Chongqing, 400037, PR China.

20 E-mail address: xuyali1976@163.com

21

22 **Abstract**

23 **Background:** Due to exposure to hypoxic environments, individuals residing in plateau regions
24 are susceptible to pulmonary hypertension (PH). Consequently, there is an urgent need for a
25 simple and efficient nomogram to assess the risk of PH in this population.

26 **Methods:** This study included a total of 6,603 subjects, who were randomly divided into a
27 validation set and a derivation set at a ratio of 7:3. Optimal predictive features were identified
28 through the least absolute shrinkage and selection operator regression technique, and nomograms
29 were constructed using multivariate logistic regression. The performance of these nomograms was
30 evaluated and validated using the area under the curve (AUC), calibration curves, the
31 Hosmer-Lemeshow test, and decision curve analysis. Comparisons between nomograms were
32 conducted using the net reclassification improvement (NRI) and integrated discrimination
33 improvement (IDI) indices.

34 **Results:** Nomogram^I was established based on independent risk factors, including gender, Tibetan
35 ethnicity, age, incomplete right bundle branch block (IRBBB), atrial fibrillation (AF), sinus
36 tachycardia (ST), and T wave changes (TC). The AUCs for Nomogram^I were 0.716 in the
37 derivation set and 0.718 in the validation set. Nomogram^{II} was established based on independent
38 risk factors, including Tibetan ethnicity, age, right axis deviation (RAD), high voltage in the right
39 ventricle (HVRV), IRBBB, AF, pulmonary P waves, ST, and TC. The AUCs for Nomogram^{II} were
40 0.844 in the derivation set and 0.801 in the validation set. Both nomograms demonstrated
41 satisfactory clinical consistency. The IDI and NRI indices confirmed that Nomogram^{II}
42 outperformed Nomogram^I. Therefore, the online dynamic Nomogram^{II} was established.

43 **Conclusions:** A reliable and straightforward nomogram was developed to predict the risks of PH
44 in the plateau population.

45

46 **Introduction**

47 Pulmonary hypertension (PH) is a chronic, progressive condition characterised by elevated
48 pulmonary arterial pressure, primarily resulting from pulmonary vascular remodelling. This
49 remodelling is driven by the infiltration of inflammatory cells, endothelial-to-mesenchymal
50 transition, and hyperplasia of the pulmonary intima (Rubin & Naeije, 2023; Shah, Beckmann,
51 Vorla, & Kalra, 2023; Simonneau et al., 2019). PH often presents similarly to other lung diseases,
52 leading to diagnostic delays and, consequently, delays in receiving optimal treatment.
53 Approximately 1% of the adult population and more than half of individuals with congestive heart
54 failure are affected by PH (Hoepfer et al., 2016; Mandras, Mehta, & Vaidya, 2020). Moreover, as
55 the pulmonary vascular load increases, PH can ultimately lead to life-threatening right heart
56 failure. The 1-year and 3-year survival rates for patients with PH range from 68% to 93% and 39%
57 to 77%, respectively (Naeije, Richter, & Rubin, 2022; Ruopp & Cockrill, 2022).

58 Right-heart catheterisation (RHC) is recognised as the gold standard for diagnosing PH,
59 clarifying the specific diagnosis, and determining the severity of the condition. However, due to its
60 invasive nature, RHC is not suitable as a widespread population screening tool for PH (McGoon et
61 al., 2004). Transthoracic echocardiography (TTE), a non-invasive screening test, is extensively
62 used for PH because it can provide estimates of pulmonary arterial systolic pressure (sPAP) and
63 evaluates cardiac structure and function. A clinical study involving 731 patients in China found no
64 significant difference between RHC and TTE in assessing sPAP in PH caused by hypoxia.

65 Furthermore, Pearson correlation analysis between RHC and TTE demonstrated a moderate
66 overall correlation (Hong et al., 2023; McGoon et al., 2004; Xu & Jing, 2009).

67 According to literature reviews, nearly 140 million individuals reside in high-altitude regions
68 (altitudes exceeding 2,500 meters), and the number of people visiting these areas for economic or
69 recreational reasons has been increasing over the past few decades (Moore, Niermeyer, &
70 Zamudio, 1998; West, 2012; Xu & Jing, 2009). High altitude typically signifies a hypoxic
71 environment due to the decrease in barometric pressure as altitude increases, which proportionally
72 reduces PO₂, resulting in hypobaric hypoxia (Gassmann et al., 2021). PH arising from prolonged
73 exposure to hypoxic conditions at high altitudes is termed high-altitude pulmonary hypertension
74 (Xu & Jing, 2009). Hypoxia triggers hypoxic pulmonary vasoconstriction (HPV), a physiological
75 response aimed at optimising ventilation/perfusion matching by redirecting blood to
76 better-oxygenated segments of the lung through the constriction of small pulmonary arteries
77 (Dunham-Snary et al., 2017). Furthermore, sustained hypoxia leads to pulmonary vascular
78 remodelling, increasing resistance to blood flow due to reduced vessel elasticity and decreased
79 vessel diameter. HPV and vascular remodelling are the primary mechanisms underlying
80 hypoxia-induced PH, which significantly impairs right ventricular function and can ultimately
81 result in fatal heart failure (Julian & Moore, 2019; Penaloza & Arias-Stella, 2007). Consequently,
82 there is a pressing need for a straightforward and dependable model to assist clinicians and
83 individuals in assessing the risk of PH in populations at high altitudes.

84 In this study, we developed and validated two risk prediction models for high-altitude PH
85 based on TTE results by examining routine inspection parameters in Tibet, China.

86

87 **Materials and methods**

88 *Study population and data collection*

89 Upon gathering data from all patients who underwent both TTE and 12-lead electrocardiogram
90 (ECG) examinations at the General Hospital of Tibet Military Area Command between April 2021
91 and October 2023, we further screened the records based on the following criteria: (1) age > 14
92 years; (2) interval between the TTE and ECG examinations < 2 months, and (3) for patients with
93 multiple TTE and/or ECG records, only the examination with the shortest interval between TTE
94 and ECG was selected. Ultimately, we compiled examination data for 6,603 eligible patients.

95 The retrospectively-collected clinical data were categorised into two main groups: (1)
96 demographic characteristics, including name, age, gender, and Tibetan ethnicity; (2) ECG results,
97 encompassing right axis deviation (RAD), clockwise rotation (CR), counterclockwise rotation
98 (CCR), high voltage in the right ventricle (HVRV), incomplete right bundle branch block
99 (IRBBB), complete right bundle branch block (CRBBB), atrial fibrillation (AF), sinus arrhythmia
100 (SA), sinus bradycardia (SB), sinus tachycardia (ST), T wave changes (TC), ST-segment changes
101 (STC), atrial premature beats (APB), ventricular premature beats (VPB), junctional premature
102 beats (JPB), complete left bundle branch block (CLBBB), first-degree atrioventricular block
103 (IAB), and pulmonary P waves (PP); (3) TTE results: pulmonary arterial systolic pressure (sPAP)
104 was measured via TTE to evaluate PH. PH was graded as follows: Grade I PH ($50 \text{ mmHg} > \text{sPAP}$
105 $\geq 30 \text{ mmHg}$), Grade II PH ($70 \text{ mmHg} > \text{sPAP} \geq 50 \text{ mmHg}$), and Grade III PH ($\text{sPAP} \geq 70$
106 mmHg). The severity of PH increases with its grade, indicating a higher risk of the condition.

107 All procedures were conducted following the approval of the Ethics Committee of the General
108 Hospital of Tibet Military Area Command (APPROVAL NUMBER: 2024-KD002-01).

109 Subsequently, the data from all participants were anonymised and de-identified prior to analysis.

110 Consequently, the requirement for informed consent was waived.

111

112 *Statistical analysis*

113

114 Statistical analysis was performed with R software version 4.3.2. $P < 0.05$ (double-tailed)

115 was considered statistically significant.

116 For validation and derivation of the prediction model, subjects were divided into a

117 validation set and a derivation set randomly, at a ratio of 7:3, respectively. Categorical variables

118 were transformed into dichotomous variables, and continuous variables were expressed by

119 concrete values (means \pm standard deviation) and analysed using Student's t-test. Fisher's exact

120 test or Pearson's χ^2 test was applied for categorical variables.

121 The derivation set was used to select optimal predictive factors through the least absolute

122 shrinkage and selection operator (LASSO) regression technique. Independent factors were

123 identified via multivariate logistic regression analysis, incorporating variables selected during the

124 LASSO regression. A backward step-down selection process, guided by the Akaike information

125 criterion, determined the final model. The predictive accuracy of the nomograms was assessed

126 using the AUC of the ROC curve in both the derivation and validation sets. The

127 Hosmer-Lemeshow test and calibration curves were employed to evaluate the consistency between

128 actual outcomes and predicted probabilities. The clinical utility of the nomograms was assessed

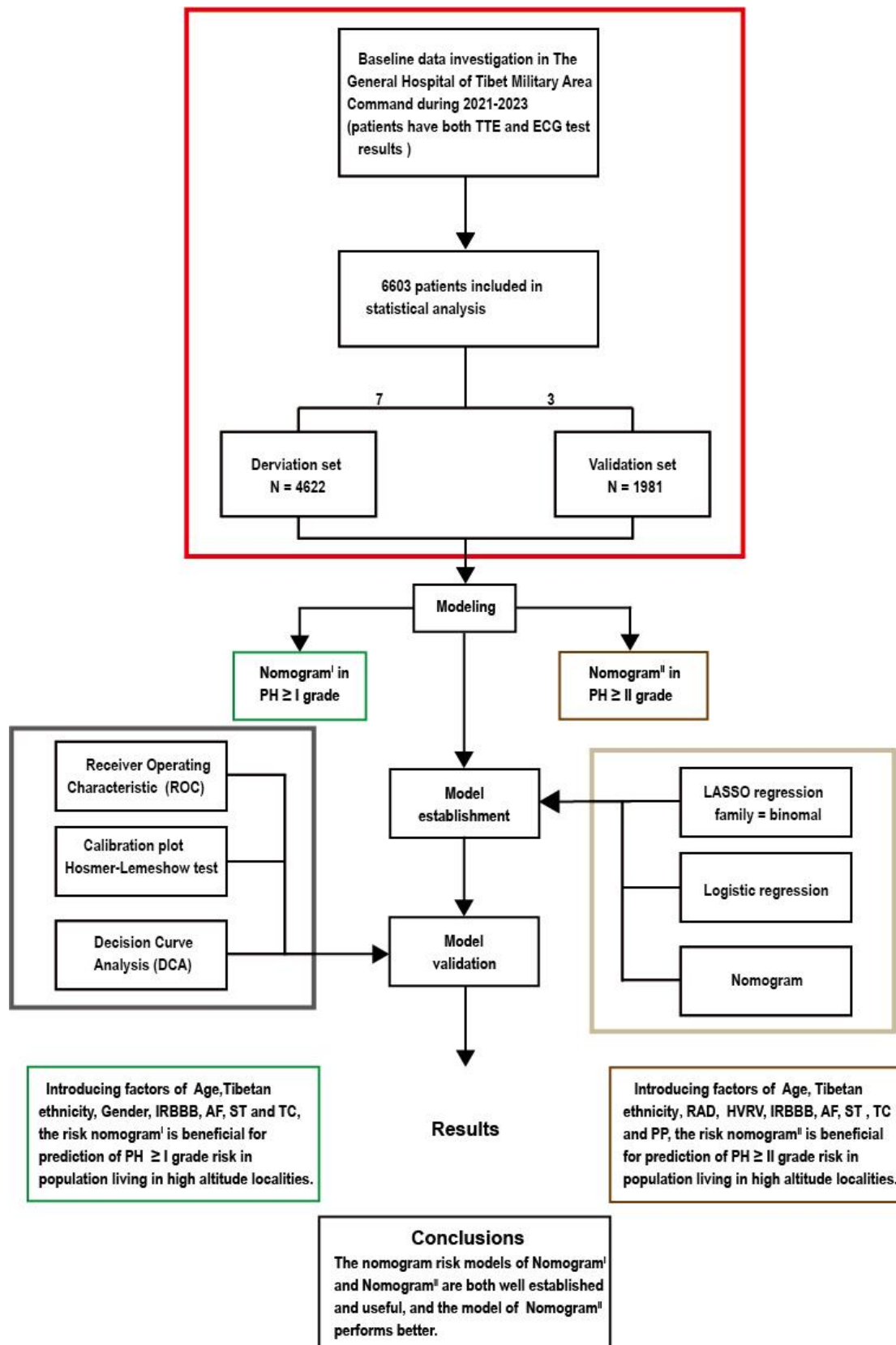
129 through decision curve analysis (DCA). The cut-off value for the total score in the nomogram was

130 established based on the ROC curve, with patients categorised into low-risk and high-risk groups.

131 The performance comparison between nomograms was analysed using the integrated

132 discrimination improvement (IDI) and net reclassification improvement (NRI).

133



134

135

136 Figure 1. Flow diagram. Based on the exclusion and inclusion criteria, 6,603 patients were
137 included in this study. Patients were divided into a validation set and a derivation set randomly
138 following a 7:3 ratio. pulmonary hypertension, PH; right axis deviation, RAD; high voltage in the
139 right ventricle, HVRV; incomplete right bundle branch block, IRBBB; atrial fibrillation, AF; sinus
140 tachycardia, ST; T wave changes, TC; Pulmonary P waves, PP.

141

142 Results

143

144 **Table 1. Baseline characteristics of individuals in the derivation and validation**
145 **sets**

Variable	Derivation set (n = 4622)	Validation set (n = 1981)	P
Age			
Total (Mean ± SD)	42.43 ± 16.93	42.05 ± 16.41	0.390
Age ≤ 42, n (%)	2619(56.66)	1135(57.29)	
Age > 42, n (%)	2003(43.34)	846(42.71)	0.635
Tibetan, n (%)			0.538
No	2856 (61.79)	1240 (62.59)	
Yes	1766 (38.21)	741 (37.41)	
Gender, n (%)			0.260
female	1219 (26.37)	549 (27.71)	

male	3403 (73.63)	1432 (72.29)	
RAD, n (%)			0.141
No	3833 (82.93)	1672 (84.40)	
Yes	789 (17.07)	309 (15.60)	
CR, n (%)			0.387
No	4000 (86.54)	1730 (87.33)	
Yes	622 (13.46)	251 (12.67)	
CCR, n (%)			0.402
No	3994 (86.41)	1727 (87.18)	
Yes	628 (13.59)	254 (12.82)	
HVRV, n (%)			0.102
No	4151 (89.81)	1805 (91.12)	
Yes	471 (10.19)	176 (8.88)	
IRBBB, n (%)			0.573
No	4547 (98.38)	1945 (98.18)	
Yes	75 (1.62)	36 (1.82)	
CRBBB, n (%)			0.945
No	4444 (96.15)	1904 (96.11)	
Yes	178 (3.85)	77 (3.89)	
AF, n (%)			0.594

No	4551 (98.46)	1954 (98.64)	
Yes	71 (1.54)	27 (1.36)	
SA, n (%)			0.243
No	4247 (91.89)	1837 (92.73)	
Yes	375 (8.11)	144 (7.27)	
ST, n (%)			0.910
No	4395 (95.09)	1885 (95.15)	
Yes	227 (4.91)	96 (4.85)	
SB, n (%)			0.345
No	4245 (91.84)	1833 (92.53)	
Yes	377 (8.16)	148 (7.47)	
TC, n (%)			0.769
No	4003 (86.61)	1721 (86.88)	
Yes	619 (13.39)	260 (13.12)	
STC, n (%)			0.415
No	4399 (95.18)	1876 (94.70)	
Yes	223 (4.82)	105 (5.30)	
APB, n (%)			0.219
No	4587 (99.24)	1960 (98.94)	
Yes	35 (0.76)	21 (1.06)	

JPB, n (%)			0.425
No	4603 (99.59)	1970 (99.44)	
Yes	19 (0.41)	11 (0.56)	
VPB, n (%)			0.844
No	4580 (99.09)	1962 (99.04)	
Yes	42 (0.91)	19 (0.96)	
PP, n (%)			0.439
No	4507 (97.51)	1938 (97.83)	
Yes	115 (2.49)	43 (2.17)	
CLBBB, n (%)			0.757
No	4610 (99.74)	1975 (99.70)	
Yes	12 (0.26)	6 (0.30)	
IAB, n (%)			0.910
No	4556 (98.57)	1952 (98.54)	
Yes	66 (1.43)	29 (1.46)	
PH \geq I grade, n (%)			0.820
No	2793 (60.43)	1203 (60.73)	
Yes	1829 (39.57)	778 (39.27)	
PH \geq II grade, n(%)			0.962
No	4227 (91.45)	1811 (91.42)	

Yes

395 (8.55)

170 (8.58)

146

147 *Subjects' characteristics*

148

149 Following a 7:3 allocation ratio, 4,622 subjects were placed in the derivation set and 1,981
150 subjects in the validation set. The characteristics of the subjects are presented in Table 1. The
151 prevalence of PH of Grade I or higher was 39.57% (1,829 cases) in the derivation set and 39.27%
152 (778 cases) in the validation set ($P=0.820 > 0.05$). The prevalence of PH of Grade II or higher was
153 8.55% (395 cases) in the derivation set and 8.58% (170 cases) in the validation set ($P=0.962 >$
154 0.05). No significant difference was observed in the age distribution between the derivation and
155 validation sets (42.43 ± 16.93 vs. 42.05 ± 16.41 , $P=0.390 > 0.05$), with age categorised into ≤ 42
156 and > 42 subgroups based on the mean age. The composition ratios of the two age subgroups did
157 not significantly differ between the validation and derivation sets ($P=0.6352 > 0.05$). Furthermore,
158 no significant differences were observed in the characteristics related to gender, Tibetan or not,
159 RAD, CR, CCR, HVRV, IRBBB, CRBBB, AF, SA, ST, SB, TC, STC, APB, VPB, JPB, PP, IAB,
160 and CLBBB. (Table 1)

161

162 *Independent risk factors in $PH \geq I$ grade group and $PH \geq II$ grade group*

163

164 In the $PH \geq I$ grade group, based on the λ_{\min} criterion in the LASSO regression model, 18
165 out of 22 variables were selected. However, this selection was deemed excessive for practical
166 clinical applications. Therefore, we further refined the model using the λ_{1se} criterion, which
167 reduced the number of variables, albeit with a significant decrease in the AUC of the ROC curve
168 (λ_{1se}) compared to the ROC curve (λ_{\min}) (Fig. 2 C, E, G). Ultimately, 9 variables were chosen

169 according to λ_{1se} , including gender, Tibetan ethnicity, age ≤ 42 , age > 42 , IRBBB, CRBBB, AF,
 170 ST, and TC (Fig. 2 I). Gender, Tibetan ethnicity, age, IRBBB, AF, ST, and TC were subsequently
 171 identified as independent risk factors for PH \geq I grade through multivariate logistic regression
 172 analysis and were used to develop Nomogram^I. (Table 2)

173 In the PH \geq II grade group, based on the λ_{1se} criterion in the LASSO regression model (Fig.
 174 2 D, F), 11 variables were selected to align with clinical needs. These variables were Tibetan
 175 ethnicity, age ≤ 42 , age > 42 , RAD, HVRV, IRBBB, CRBBB, AF, PP, ST, and TC (Fig. 2 J).
 176 Tibetan ethnicity, age, RAD, HVRV, IRBBB, AF, PP, ST, and TC were determined to be
 177 independent risk factors for PH \geq II grade through multivariate logistic regression analysis and
 178 were utilised to construct Nomogram^{II}. (Table 3).

179

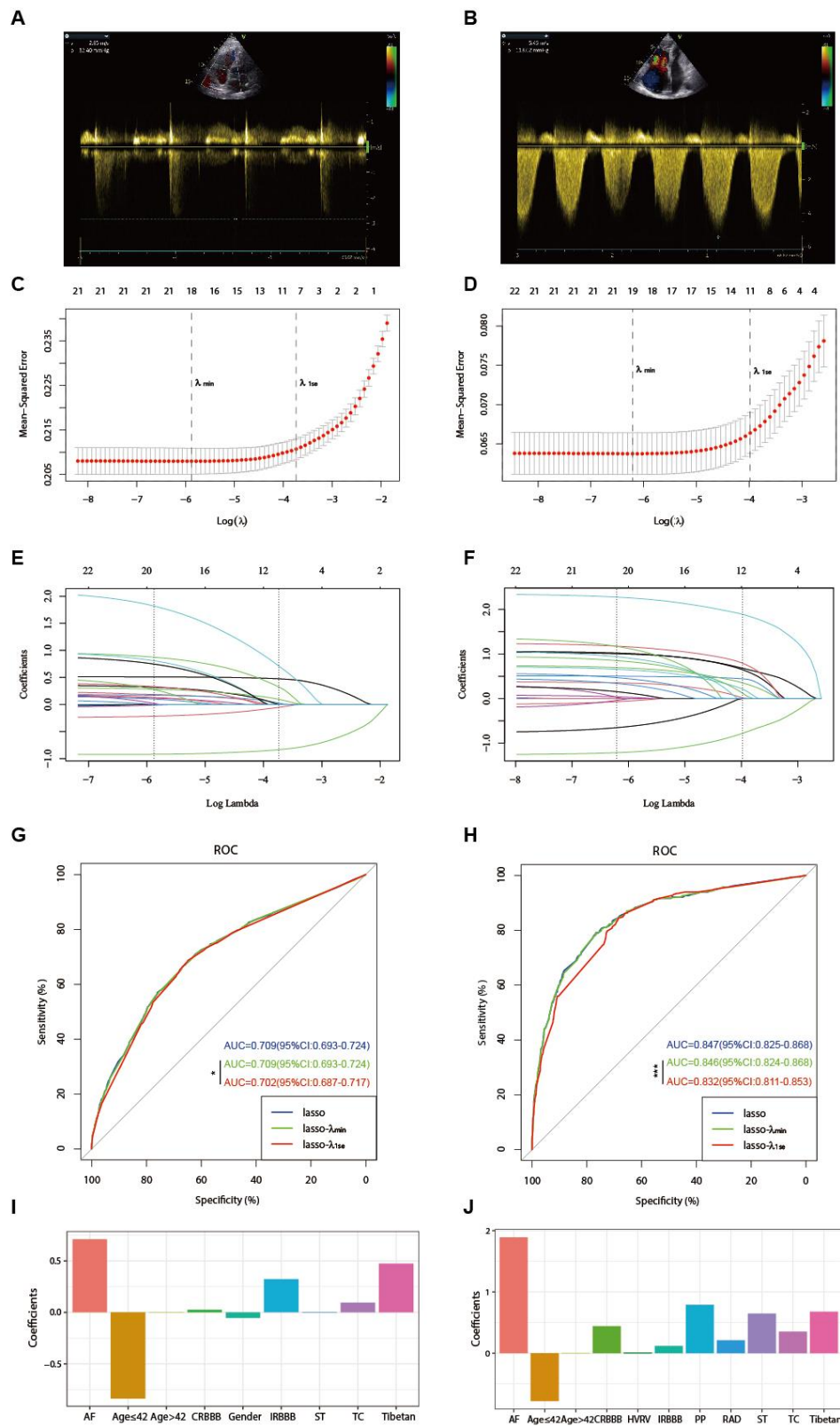
180 **Table 2. Risk factors for PH \geq I grade in the derivation set**

Variable	β -Coefficient	OR (95%CI)	P
Tibetan	0.34	1.40 (1.23-1.60)	<0.001
Gender	-0.3	0.74 (0.65-0.84)	<0.001
Age	0.034	1.03 (1.03-1.04)	<0.001
IRBBB	1.106	3.02 (1.96-4.67)	<0.001
AF	1.431	4.18 (2.19-7.97)	<0.001
ST	0.369	1.45 (1.14-1.84)	0.003
TC	0.306	1.36 (1.16-1.59)	<0.001

181

Table 3. Risk factors for PH \geq II grade in the derivation set

Variable	β -Coefficient	OR (95%CI)	P
Tibetan	0.689	1.99 (1.55-2.57)	<0.001
Age	0.042	1.04 (1.03-1.05)	<0.001
RAD	0.751	2.12 (1.56-2.88)	<0.001
HVRV	0.486	1.63 (1.14-2.31)	0.007
IRBBB	1.512	4.53 (2.77-7.42)	<0.001
AF	2.102	8.18 (5.13-13.05)	<0.001
ST	1.247	3.48 (2.58-4.70)	<0.001
TC	0.592	1.81 (1.44-2.27)	<0.001
PP	1.486	4.42 (2.96-6.61)	<0.001



184

185

186

Figure 2 illustrates the optimal predictive variables as determined by the LASSO binary

187

logistic regression model. Panels A and B depict the measurement of tricuspid regurgitation

188 spectra via transthoracic echocardiography in patients with Grade I PH (A) and Grade III PH (B).
189 Panels C to J demonstrate the identification of the optimal penalisation coefficient lambda (λ) in
190 the LASSO model using 10-fold cross-validation for the PH \geq I grade group (C) and the PH \geq II
191 grade group (D). The dotted line on the left (λ_{\min}) represents the value of the harmonic
192 parameter $\log(\lambda)$ at which the model's error is minimised, and the dotted line on the right (λ_{1se})
193 indicates the value of the harmonic parameter $\log(\lambda)$ at which the model's error is minimal minus 1
194 standard deviation. The LASSO coefficient profiles of 22 predictive factors for the PH \geq I grade
195 group (E) and the PH \geq II grade group (F) show that as the value of λ decreased, the degree of
196 model compression increased, enhancing the model's ability to select significant variables. ROC
197 curves were constructed for three models (LASSO, LASSO- λ_{\min} , LASSO- λ_{1se}) in both the PH
198 \geq I grade group (G) and the PH \geq II grade group (H). Histograms depict the final variables selected
199 according to λ_{1se} and their coefficients for the PH \geq I grade group (I) and the PH \geq II grade group
200 (J). Asterisks denote levels of statistical significance: *P < 0.05, **P < 0.01, ***P < 0.001.

201

202 *Construction of Nomogram^I in PH \geq I grade group and Nomogram^{II} in PH \geq II grade group*

203

204 In the PH \geq I grade group, a predictive Nomogram^I for PH \geq I grade was developed based on
205 independent risk factors, including gender, Tibetan ethnicity, age, IRBBB, AF, ST, and TC. Points
206 are assigned to each independent factor by drawing a vertical line to the points scale. The total
207 points for an individual correspond to their risk of developing PH. Patients were then classified
208 into high-risk and low-risk subgroups according to the total score's cut-off value (cut-off value:
209 45), which was determined based on the ROC curve (Fig. 3 A). The risks for the two groups were
210 evaluated in both the derivation and validation sets. In the derivation set, the risk of PH in the

211 high-risk group was significantly higher than in the low-risk group (odds ratio [OR]: 4.210, 95%
212 confidence interval [CI]: 3.715-4.775) (Fig. 3 B), as was also observed in the validation set (odds
213 ratio [OR]: 4.207, 95% confidence interval [CI]: 3.476-5.102) (Fig. 3 C).

214 In the PH \geq II grade group, a predictive Nomogram^{II} for PH \geq II grade was developed using
215 independent risk factors, including Tibetan ethnicity, age, RAD, HVRV, IRBBB, AF, PP, ST and
216 TC. Based on the cut-off value of the total score (cut-off value: 76), determined in line with the
217 ROC curve, patients were categorised into high-risk and low-risk subgroups (Fig. 3 D). The risks
218 for the two groups were evaluated in both the derivation and validation sets. In the derivation set,
219 the risk of PH in the high-risk group was significantly greater than in the low-risk group (odds
220 ratio [OR]: 11.591, 95% confidence interval [CI]: 9.128-14.845) (Fig. 3 E), a finding that was
221 replicated in the validation set (odds ratio [OR]: 7.103, 95% confidence interval [CI]: 5.106-9.966)
222 (Fig. 3 F)

223 .

224



225

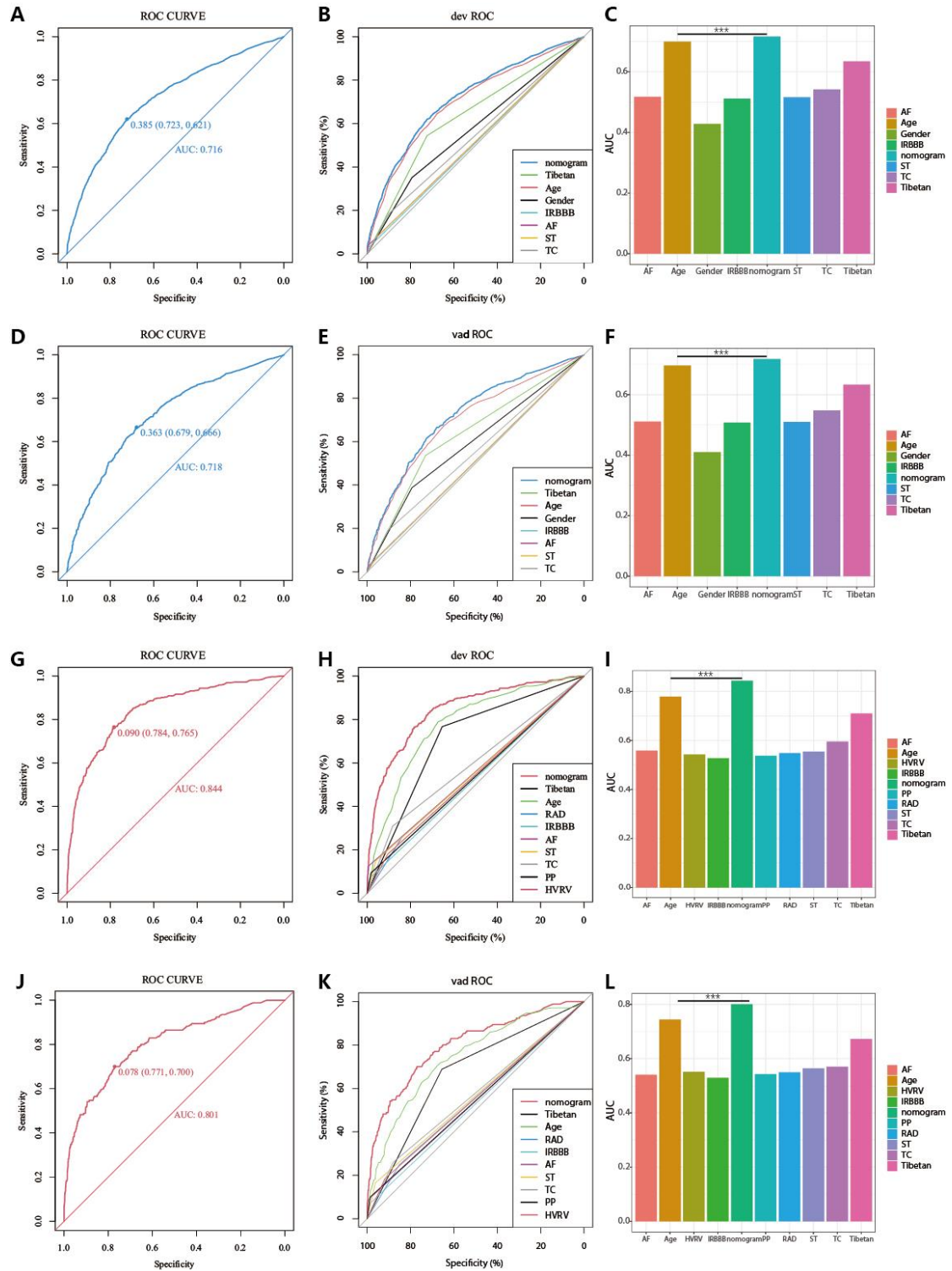
226 Figure 3. Nomogram for predicting PH and risk stratification based on total score. (A-C)

227 Nomogram^I for the prediction of PH \geq I grade in the PH \geq I grade group. Points for each

228 independent factor are summed to calculate total points, determining the corresponding ‘risk’ level.

229 Patients were divided into ‘High-risk’ and ‘Low-risk’ subgroups according to the cutoff of the total

230 points (A). Histograms illustrate the odds ratio (OR) comparing the 'High-risk' group to the
231 'Low-risk' group in the derivation set (B) and validation set (C). (D-F) Nomogram^{II} for predicting
232 PH \geq II grade within the PH \geq II grade group: Similarly, points from each independent factor are
233 totalled, and the corresponding 'risk' level is ascertained. Patients are divided into 'High-risk' and
234 'Low-risk' groups based on the cut-off value of the total points (D). Histograms display the OR for
235 the 'High-risk' group compared to the 'Low-risk' group in the derivation (E) and validation set (F).
236 *** P < 0.001. (G) Screenshot of dynamic Nomogram^{II}'s web page.
237



238

239 Figure 4. Receiver operating characteristic (ROC) curves and area under the curve (AUC)

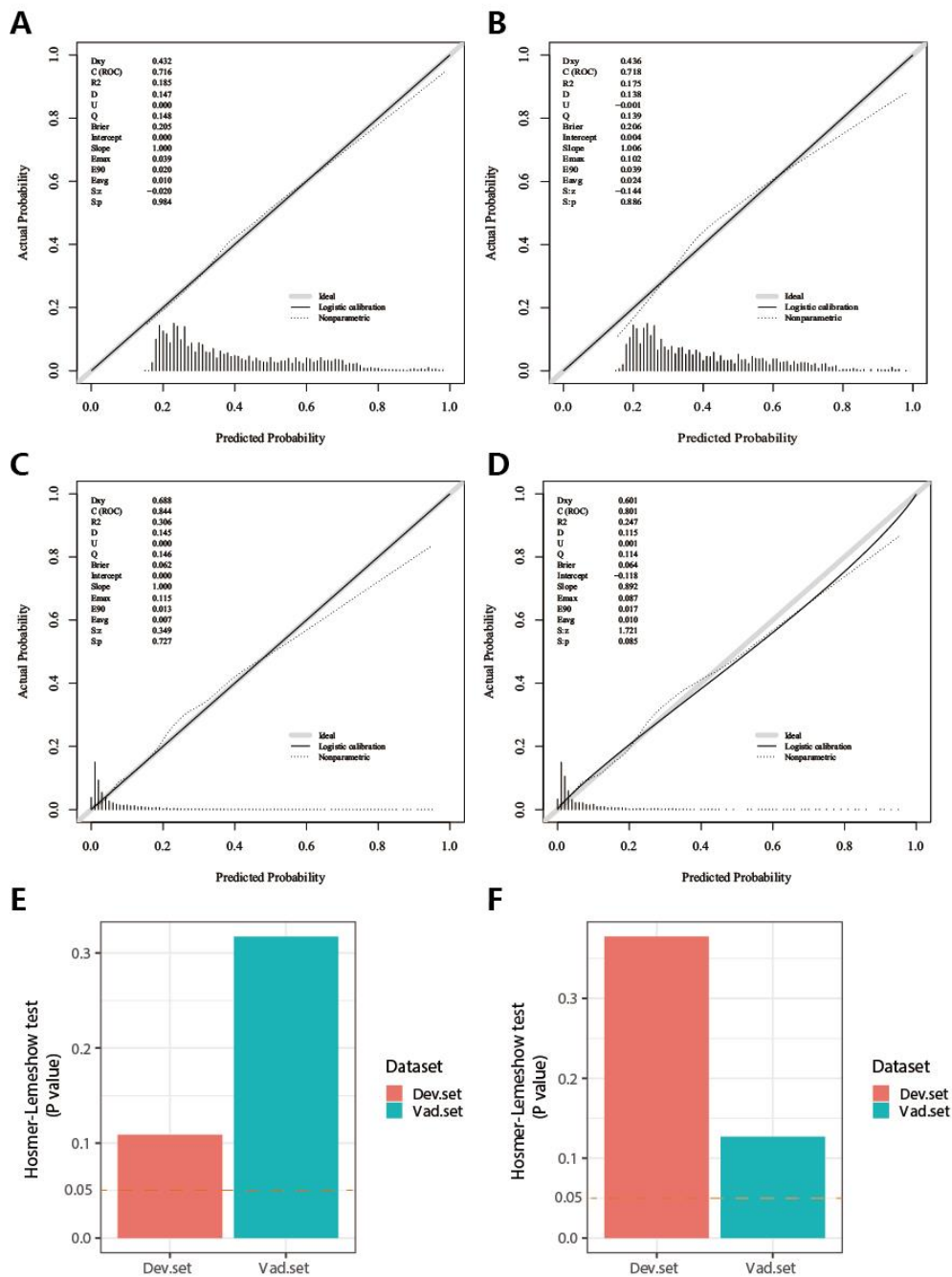
240 for Nomogram^I in PH \geq I and Nomogram^{II} in PH \geq II grade groups. (A-F) In the PH \geq I grade group,

241 the ROC and corresponding AUC of Nomogram^I and independent factors in the derivation set

242 (A-C) and validation set (D-F). (G-L) In the PH \geq II grade group, the ROC and corresponding AUC

243 of Nomogram^{II} and independent factors in the derivation set (G-I) and validation set (J-L).

244



245

246 Figure 5. Calibration plots and Hosmer-Lemeshow test results for Nomogram^I in PH ≥ I and

247 Nomogram^{II} in PH ≥ II grade groups. (A-B) In the PH ≥ I grade group, the calibration plots of

248 Nomogram^I in the derivation set (A) and the validation set (B). (C-D) In the PH ≥ II grade group,

249 the calibration plots of Nomogram^{II} in the derivation set (C) and the validation set (D). (E) In the

250 PH \geq I grade group, Hosmer-Lemeshow test results for Nomogram^I in the derivation set and the
251 validation set. (F) In the PH \geq II grade group, Hosmer-Lemeshow test results for Nomogram^{II} in the
252 derivation set and the validation set.

253

254 *Assessment and validation of Nomogram^I in the PH \geq I grade group and Nomogram^{II} in the*
255 *PH \geq II grade group*

256

257 In the PH \geq I grade group, Nomogram^I was developed to predict the risk of PH \geq I grade,
258 utilising the AUC to assess its discriminative ability. The AUC value for Nomogram^I was 0.716
259 (95% confidence interval [CI]: 0.701 - 0.731) in the derivation set (Fig 4. A) and 0.718 (95%
260 confidence interval [CI]: 0.695 - 0.741) in the validation set (Fig 4. D). Furthermore, ROC curves
261 were used to compare the discriminative capacity of Nomogram^I and single independent factors in
262 predicting PH \geq I grade. Notably, the AUC of Nomogram^I was significantly higher than that of any
263 single independent factor in the derivation (Fig 4. B, C) and the validation set (Fig 4. E, F). The
264 calibration curves for the derivation set (Fig 5. A) and the validation set (Fig 5. B) demonstrated
265 high agreement between predicted and actual values, indicating that Nomogram^I accurately
266 predicts PH \geq I grade. The results of the Hosmer-Lemeshow test in both the derivation set
267 (P=0.109 > 0.05) and the validation set (P=0.317 > 0.05) further confirmed the effective
268 performance of Nomogram^I (Fig 5. E).

269 Nomogram^{II} was developed to predict the risk of PH \geq II grade. The AUC for Nomogram^{II}
270 was 0.844 (95% confidence interval [CI]: 0.823 - 0.865) in the derivation set (Fig 4. G) and 0.801
271 (95% confidence interval [CI]: 0.763 - 0.838) in the validation set (Fig 4. J). Furthermore, ROC
272 curves were used to compare the discriminative capacity of Nomogram^{II} and individual

273 independent factors in predicting PH \geq II grade. The AUC of Nomogram^{II} was significantly higher
274 than that of any single independent factor in the derivation set (Fig 4. H, I) and the validation set
275 (Fig 4. K, L). The calibration curves for the derivation set (Fig 5. C) and the validation set (Fig 5.
276 D) demonstrated high agreement between the predicted and actual values, indicating that
277 Nomogram^{II} accurately predicts PH \geq II grade. Additionally, the results of the Hosmer-Lemeshow
278 test in the derivation set (P=0.377 > 0.05) and the validation set (P=0.127 > 0.05) further
279 confirmed the good performance of Nomogram^{II} (Fig 5. F).

280
281
282

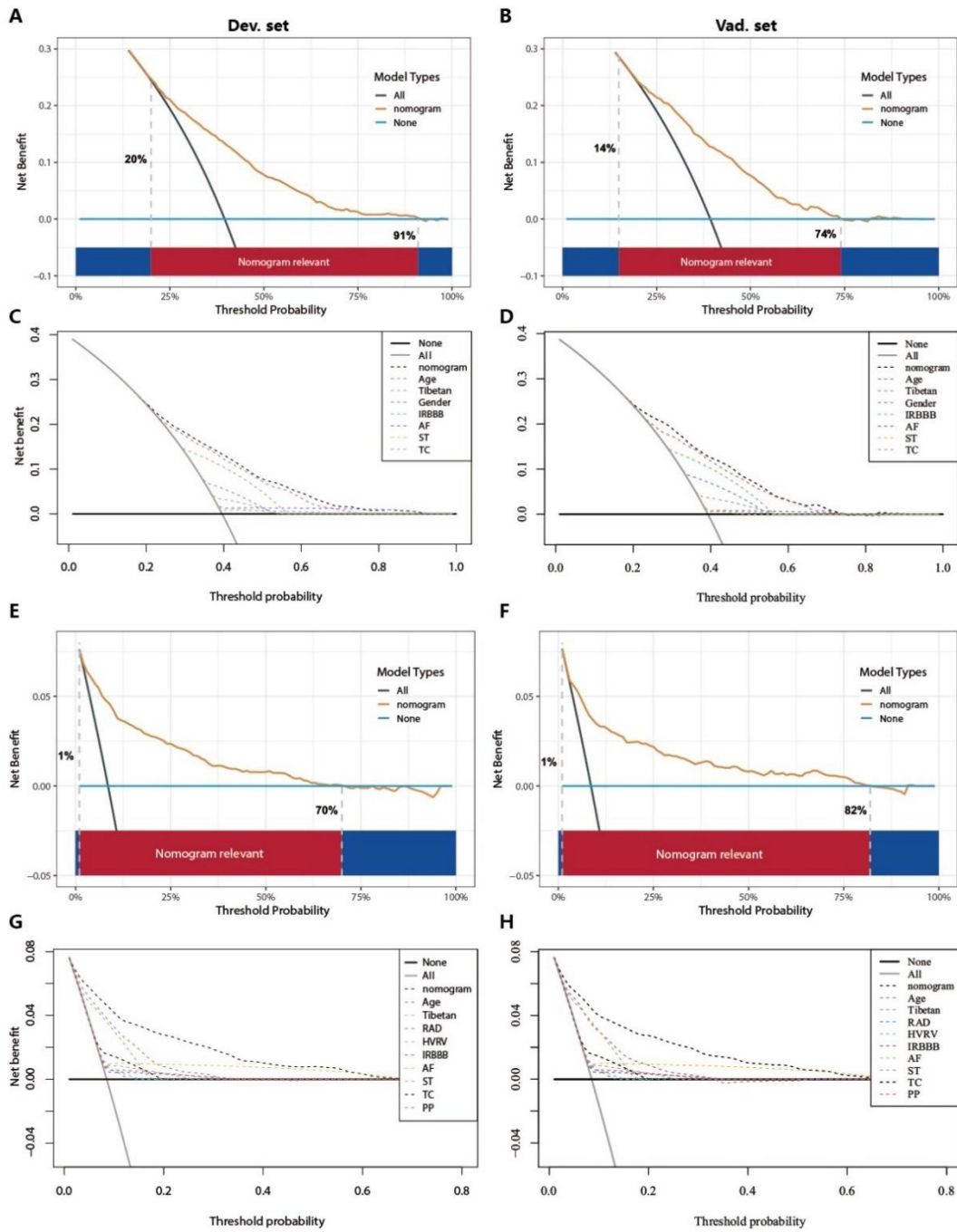
Clinical utility of Nomogram^I and Nomogram^{II}

283 In the PH \geq I grade group, the clinical utility of Nomogram^I for predicting the risk of PH \geq I
284 grade was assessed using DCA. This analysis revealed a significant net benefit with a threshold
285 probability range of 20% to 91% in the derivation set (Fig 6. A) and 14% to 74% in the validation
286 set (Fig 6. B). Moreover, the DCA curve from the derivation set indicated that the clinical
287 predictive capability of Nomogram^I surpassed that of any single independent factor, a finding that
288 was corroborated in the validation set (Fig 6. C, D).

289

290 In the PH \geq II grade group, the clinical utility of Nomogram^{II} for predicting the risk of PH \geq II
291 grade was evaluated using DCA, which showed a clear net benefit within the threshold probability
292 range of 1% to 70% in the derivation set (Fig 6. E) and 1% to 82% in the validation set (Fig 6. F).
293 Additionally, the DCA curve for the derivation set demonstrated that the clinical predictive
294 effectiveness of Nomogram^{II} exceeded that of any single independent factor, a conclusion that was
295 also confirmed in the validation set (Fig 6. G, H).

296



298

299 Figure 6. Decision curve analysis (DCA) for Nomogram^I in the PH≥I grade and Nomogram^{II}

300 in the PH≥II grade group. (A-D) In the PH≥I grade group, the DCAs of Nomogram^I and

301 independent factors in the derivation (A, C) and validation set (B, D). (E-H) In the PH≥II grade

302 group, the DCAs of Nomogram^{II} and independent factors in the derivation (E, G) and validation

303 set (F, H).

304

305 *Comparison between Nomogram^I and Nomogram^{II}*

306 In the PH \geq I grade group, when comparing Nomogram^I to Nomogram^{II}, Nomogram^I exhibited
307 an IDI of -0.0012 (95% CI: -0.0032 to 0.0009, p=0.2777), a categorical NRI of 0.0117 (95% CI:
308 -0.0004 to 0.0237, p=0.0575), and a continuous NRI of -0.2423 (95% CI: -0.2992 to -0.1854,
309 p<0.001) in predicting the risk of PH \geq I grade.

310 In the PH \geq II grade group, compared to Nomogram^I, Nomogram^{II} demonstrated an IDI of
311 0.0366 (95% CI: 0.0247 to 0.0485, p<0.001), a categorical NRI of 0.0301 (95% CI: 0.0093 to
312 0.0510, p<0.05), and a continuous NRI of 0.2785 (95% CI: 0.1824 to 0.3745, p<0.001) for
313 predicting the risk of PH \geq II grade.

314 These results indicate that Nomogram^{II} outperformed Nomogram^I in terms of IDI and NRI
315 values.

316 *Website of Nomogram^{II}*

317 Patients and physicians can calculate the risk of pulmonary hypertension through a free
318 web-based dynamic Nomogram^{II} (<https://dapeng.shinyapps.io/dynnomapp-1/>).

319

320 **Discussion**

321 A significant portion of the global population lives in high-altitude areas such as the Tibetan
322 Plateau, Ethiopian Highlands, Andes Mountains, and Pamir Plateau. These regions are marked by
323 an extremely hypoxic environment that leads to alveolar hypoxia, posing severe risks to the
324 cardiopulmonary system. One such risk is the development of PH, which occurs through
325 mechanisms like hypoxic pulmonary vasoconstriction and pulmonary vascular remodelling

326 (Burtscher, Gatterer, Burtscher, & Mairböurl, 2018; Sydykov et al., 2021; Wilkins, Ghofrani,
327 Weissmann, Aldashev, & Zhao, 2015). Accurate, timely diagnosis and early, effective treatment
328 are crucial for the clinical improvement and survival of patients with PH. Without prompt
329 intervention, PH can impair right heart function and ultimately result in fatal right heart failure
330 (Benza et al., 2010; Kim & George, 2019; McGoon et al., 2004). Thus, there is a need to develop a
331 predictive model to estimate the risk of PH, facilitating risk stratification and management. In this
332 study, we analysed routine electrocardiogram examination indicators and basic demographic
333 information to assess the risk of PH. We developed nomograms for the PH \geq I grade group and the
334 PH \geq II grade group, and the performance of these nomograms was evaluated and validated.

335 Currently, TTE is widely utilised for large-scale, non-invasive screening of patients at risk for
336 PH (D'Alto et al., 2018; Habib & Torbicki, 2010; Janda, Shahidi, Gin, & Swiston, 2011). However,
337 in plateau regions such as Tibet, medical resources are relatively limited, and remote villages and
338 towns lack the facilities for TTE examinations. ECG examinations, being easy to administer,
339 cost-effective, and feasible for remote delivery through telemedicine, offer a practical alternative
340 (Ismail, Jovanovic, Ramzan, & Rabah, 2023). A retrospective analysis has demonstrated that ECG
341 examination results correlate with clinical parameters reflecting the severity of PH (Michalski et
342 al., 2022). Therefore, in developing this model, we primarily relied on ECG examination results
343 from patients. Utilising ECG results as predictors of PH can significantly aid clinicians in
344 identifying potential PH patients in remote plateau areas, facilitating their access to timely and
345 relevant treatment.

346 In this study, based on sPAP assessed by TTE examination, patients at risk of PH were
347 classified into grades I-III. We developed and validated two nomograms for the PH \geq I grade group

348 (Nomogram^I) and the PH \geq II grade group (Nomogram^{II}), with ECG examination results serving as
349 the primary component for both. Nomogram^I included seven variables: gender, Tibetan ethnicity,
350 age, IRBBB, AF, ST, and TC. Nomogram^{II} incorporated nine variables: Tibetan ethnicity, age,
351 RAD, HVRV, IRBBB, AF, PP, ST, and TC (Fig 3. A, D). These variables are readily available
352 from routine ECG examinations. Additionally, patients were categorised into high-risk and
353 low-risk groups based on the cut-off value of the total score in the nomogram, with the OR value
354 for the high-risk group being significantly higher than that of the low-risk group (Fig 3). Therefore,
355 both nomograms offer a useful and straightforward method for in-depth evaluation, even without
356 medical professional intervention. Both Nomogram^I and Nomogram^{II} demonstrated good
357 calibration and clinical utility (Fig 5, 6), though ROC analysis revealed that the AUC for
358 Nomogram^{II} was higher than that for Nomogram^I (0.844 vs 0.716). IDI and NRI are recognised
359 indicators that describe improved accuracy in predicting binary, multi-classification, or survival
360 outcomes (Wang, Cheng, Seaberg, & Becker, 2020). In a similar vein to a 10-year retrospective
361 cohort study, which constructed two nomograms for hypertension risk prediction and compared
362 them using IDI and NRI values (Deng et al., 2021), we used IDI and NRI to evaluate the
363 performance of Nomogram^I and Nomogram^{II}. Our findings indicated no significant difference
364 between Nomogram^I and Nomogram^{II} in the PH \geq I grade group; however, Nomogram^{II} exhibited
365 superior performance compared to Nomogram^I in the PH \geq II grade group, thus demonstrating its
366 enhanced predictive capability. So, we created an online dynamic Nomogram^{II} for doctors and
367 patients to calculate the risk of PH (Fig 3. G).

368 In this study, age and Tibetan ethnicity were identified as independent predictors of PH, a
369 finding that aligns with conclusions from a single-centre, cross-sectional study among native

370 Tibetans in Sichuan Province, China (Gou et al., 2020). We hypothesise that this association may
371 be due to the longer exposure to the hypoxic environment at high altitudes experienced by older
372 individuals and Tibetans, promoting hypoxic contraction of pulmonary blood vessels and
373 subsequent pulmonary vascular remodelling, thereby leading to PH. Additionally, the occurrence
374 of AF emerged as an independent predictor of PH with the highest OR values in both nomograms
375 (Table 2, 3). PH is known to be characterised by pulmonary vascular remodelling, which can
376 induce fibrosis and excessive myocardial apoptosis, ultimately contributing to AF (Yi et al., 2023),
377 a finding that corroborates our results. Nonetheless, it was observed that no single predictor alone
378 was effective in distinguishing PH, exhibiting poor clinical utility compared to the comprehensive
379 approach offered by the nomogram (Fig 4, Fig 6).

380 Our study has several limitations. Firstly, TTE serves only as a screening method for PH and
381 is not the gold standard; its results merely indicate the risk of PH in the examined individuals.
382 Secondly, given the constrained medical resources in remote areas, we primarily incorporated
383 readily ECG results and basic demographic information into the nomograms, resulting in a
384 relatively simple set of independent predictors. Lastly, the dataset for this study was exclusively
385 sourced from Tibet, China, meaning the validation of the nomograms lacks external validation
386 sets.

387 **Conclusion**

388 We have developed a reliable and straightforward nomogram to predict the risks associated
389 with PH, demonstrating satisfactory discrimination and calibration. Upon rigorous validation
390 using external datasets, the nomogram has shown clinical utility and favourable predictive
391 accuracy. It is anticipated to serve as an effective and convenient clinical tool for assessing the risk

392 of PH in populations residing at high altitudes.

393

394 **Acknowledgements**

395 This study was funded by the Talent Program of Army Medical University (No. 2019R038).

396

397 **Statements and Declarations**

398 The authors have no conflict of interest.

399

400 **Data availability statement**

401 Source data files have been provided.

402

403 **References**

404

405 Benza, R. L., Miller, D. P., Gomberg-Maitland, M., Frantz, R. P., Foreman, A. J., Coffey, C. S., . . .
406 McGoon, M. D. (2010). Predicting survival in pulmonary arterial hypertension: insights from the
407 Registry to Evaluate Early and Long-Term Pulmonary Arterial Hypertension Disease Management
408 (REVEAL). *Circulation*, 122:164-172. doi:10.1161/circulationaha.109.898122

409

410 Burtscher, M., Gatterer, H., Burtscher, J., & Mairböurl, H. (2018). Extreme Terrestrial
411 Environments: Life in Thermal Stress and Hypoxia. A Narrative Review. *Front Physiol*, 9:572.
412 doi:10.3389/fphys.2018.00572

413

414 D'Alto, M., Bossone, E., Opatowsky, A. R., Ghio, S., Rudski, L. G., & Naeije, R. (2018).
415 Strengths and weaknesses of echocardiography for the diagnosis of pulmonary hypertension. *Int J*
416 *Cardiol*, 263: 177-183. doi:10.1016/j.ijcard.2018.04.024

417

418 Deng, X., Hou, H., Wang, X., Li, Q., Li, X., Yang, Z., & Wu, H. (2021). Development and
419 validation of a nomogram to better predict hypertension based on a 10-year retrospective cohort
420 study in China. *Elife*, 10. doi:10.7554/eLife.66419

421

422 Dunham-Snary, K. J., Wu, D., Sykes, E. A., Thakrar, A., Parlow, L. R. G., Mewburn, J. D., . . .
423 Archer, S. L. (2017). Hypoxic Pulmonary Vasoconstriction: From Molecular Mechanisms to

424 Medicine. *Chest*, 151: 181-192. doi:10.1016/j.chest.2016.09.001
425
426 Gassmann, M., Cowburn, A., Gu, H., Li, J., Rodriguez, M., Babicheva, A., . . . Zhao, L. (2021).
427 Hypoxia-induced pulmonary hypertension-Utilizing experiments of nature. *Br J Pharmacol*, 178:
428 121-131. doi:10.1111/bph.15144
429
430 Gou, Q., Shi, R., Zhang, X., Meng, Q., Li, X., Rong, X., . . . Chen, X. (2020). The Prevalence and
431 Risk Factors of High-Altitude Pulmonary Hypertension Among Native Tibetans in Sichuan
432 Province, China. *High Alt Med Biol*, 21: 327-335. doi:10.1089/ham.2020.0022
433
434 Habib, G., & Torbicki, A. (2010). The role of echocardiography in the diagnosis and management
435 of patients with pulmonary hypertension. *Eur Respir Rev*, 19: 288-299.
436 doi:10.1183/09059180.00008110
437 Hoepfer, M. M., Humbert, M., Souza, R., Idrees, M., Kawut, S. M., Sliwa-Hahnle, K., . . . Gibbs, J.
438 S. (2016). A global view of pulmonary hypertension. *Lancet Respir Med*, 4: 306-322.
439 doi:10.1016/s2213-2600(15)00543-3
440
441 Hong, C., Chen, R., Hu, L., Liu, H., Lu, J., Zhuang, C., . . . Zheng, Z. (2023). Aetiological
442 distribution of pulmonary hypertension and the value of transthoracic echocardiography screening
443 in the respiratory department: A retrospective analysis from China. *Clin Respir J*, 17: 536-547.
444 doi:10.1111/crj.13623
445
446 Ismail, A. R., Jovanovic, S., Ramzan, N., & Rabah, H. (2023). ECG Classification Using an
447 Optimal Temporal Convolutional Network for Remote Health Monitoring. *Sensors (Basel)*, 23:
448 doi:10.3390/s23031697
449
450 Janda, S., Shahidi, N., Gin, K., & Swiston, J. (2011). Diagnostic accuracy of echocardiography for
451 pulmonary hypertension: a systematic review and meta-analysis. *Heart*, 97: 612-622.
452 doi:10.1136/hrt.2010.212084
453
454 Julian, C. G., & Moore, L. G. (2019). Human Genetic Adaptation to High Altitude: Evidence from
455 the Andes. *Genes (Basel)*, 10. doi:10.3390/genes10020150
456
457 Kim, D., & George, M. P. (2019). Pulmonary Hypertension. *Med Clin North Am*, 103: 413-423.
458 doi:10.1016/j.mcna.2018.12.002
459
460 Mandras, S. A., Mehta, H. S., & Vaidya, A. (2020). Pulmonary Hypertension: A Brief Guide for
461 Clinicians. *Mayo Clin Proc*, 95: 1978-1988. doi:10.1016/j.mayocp.2020.04.039
462
463 McGoon, M., Gutterman, D., Steen, V., Barst, R., McCrory, D. C., Fortin, T. A., & Loyd, J. E.
464 (2004). Screening, early detection, and diagnosis of pulmonary arterial hypertension: ACCP
465 evidence-based clinical practice guidelines. *Chest*, 126(1 Suppl), 14s-34s.
466 doi:10.1378/chest.126.1_suppl.14S
467

468 Michalski, T. A., Pszczola, J., Lisowska, A., Knapp, M., Sobkowicz, B., Kaminski, K., &
469 Ptaszynska-Kopczynska, K. (2022). ECG in the clinical and prognostic evaluation of patients with
470 pulmonary arterial hypertension: an underestimated value. *Ther Adv Respir Dis*, 16:
471 17534666221087846. doi:10.1177/17534666221087846
472

473 Moore, L. G., Niermeyer, S., & Zamudio, S. (1998). Human adaptation to high altitude: regional
474 and life-cycle perspectives. *Am J Phys Anthropol, Suppl* 27: 25-64.
475 doi:10.1002/(sici)1096-8644(1998)107:27+<25::aid-ajpa3>3.0.co;2-l
476

477 Naeije, R., Richter, M. J., & Rubin, L. J. (2022). The physiological basis of pulmonary arterial
478 hypertension. *Eur Respir J*, 59. doi:10.1183/13993003.02334-2021
479

480 Penalzoza, D., & Arias-Stella, J. (2007). The heart and pulmonary circulation at high altitudes:
481 healthy highlanders and chronic mountain sickness. *Circulation*, 115: 1132-1146.
482 doi:10.1161/circulationaha.106.624544
483

484 Rubin, L. J., & Naeije, R. (2023). Sotatercept for pulmonary arterial hypertension: something old
485 and something new. *Eur Respir J*, 61. doi:10.1183/13993003.01972-2022
486

487 Ruopp, N. F., & Cockrill, B. A. (2022). Diagnosis and Treatment of Pulmonary Arterial
488 Hypertension: A Review. *Jama*, 327: 1379-1391. doi:10.1001/jama.2022.4402
489

490 Shah, A. J., Beckmann, T., Vorla, M., & Kalra, D. K. (2023). New Drugs and Therapies in
491 Pulmonary Arterial Hypertension. *Int J Mol Sci*, 24. doi:10.3390/ijms24065850
492

493 Simonneau, G., Montani, D., Celermajer, D. S., Denton, C. P., Gatzoulis, M. A., Krowka, M., . . .
494 Souza, R. (2019). Haemodynamic definitions and updated clinical classification of pulmonary
495 hypertension. *Eur Respir J*, 53. doi:10.1183/13993003.01913-2018
496

497 Sydykov, A., Mamazhakypov, A., Maripov, A., Kosanovic, D., Weissmann, N., Ghofrani, H. A., . . .
498 Schermuly, R. T. (2021). Pulmonary Hypertension in Acute and Chronic High Altitude
499 Maladaptation Disorders. *Int J Environ Res Public Health*, 18. doi:10.3390/ijerph18041692
500

501 Wang, Z., Cheng, Y., Seaberg, E. C., & Becker, J. T. (2020). Quantifying diagnostic accuracy
502 improvement of new biomarkers for competing risk outcomes. *Biostatistics*.
503 doi:10.1093/biostatistics/kxaa048
504

505 West, J. B. (2012). High-altitude medicine. *Am J Respir Crit Care Med*, 186: 1229-1237.
506 doi:10.1164/rccm.201207-1323CI
507

508 Wilkins, M. R., Ghofrani, H. A., Weissmann, N., Aldashev, A., & Zhao, L. (2015).
509 Pathophysiology and treatment of high-altitude pulmonary vascular disease. *Circulation*, 131:
510 582-590. doi:10.1161/circulationaha.114.006977
511

512 Xu, X. Q., & Jing, Z. C. (2009). High-altitude pulmonary hypertension. *Eur Respir Rev*, 18: 13-17.
513 doi:10.1183/09059180.00011104

514

515 Yi, Y., Tianxin, Y., Zhangchi, L., Cui, Z., Weiguo, W., & Bo, Y. (2023). Pinocembrin attenuates
516 susceptibility to atrial fibrillation in rats with pulmonary arterial hypertension. *Eur J Pharmacol*,
517 960: 176169. doi:10.1016/j.ejphar.2023.176169

518