Appendix 1

Supplemental Methods

Measurement of clinical variables

Lifestyle variables

Lifestyle variables included age, body mass index (BMI), drinking, and smoking. BMI was calculated as weight in kilograms divided by the square of height in meters. Drinking status was divided into the following 2 categories: (I) non-drinkers (participants who never drink or who drink less than once a month, and the alcohol content is less than 10%); and (II) drinkers (patients who drink once or more a month, or the alcohol content is more than 10%). Smoking status was divided into the following 2 groups: (I) non-smokers (patients who have never smoked before); and (II) current smokers (patients who still smoke) or past smokers (those who used to smoke but have now quit).

Medical history:

Histories of hyperlipidemia, hypertension, and diabetes mellitus were obtained from in-person interviews. The diagnosis of hyperlipidemia was based on any one of the following: (I) low-density lipoprotein cholesterol (LDL-C) ≥4.14 mmol/L; (II) total cholesterol (TC) >6.45 mmol/L; and (III) triglyceride ≥2.26 mmol/L. The diagnosis of hypertension was based on resting blood pressure. If the resting blood pressure was >140/90 mmHg, the individual was diagnosed with hypertension. The diagnosis of diabetes mellitus was based any one of the following: (I) fasting plasma glucose ≥7.0 mmol/L; (II) plasma glucose ≥11.1 mmol/L 2 hours after the consumption of a drink containing 75 g of glucose; (III) symptoms of high blood sugar and a casual plasma glucose ≥11.1 mmol/L; and (IV) hemoglobin A1c ≥48 mmol/mol.

Physical examination

Blood pressure, brachial-ankle pulse wave velocity (baPWV), and ankle-brachial index (ABI) were measured using the Vascular Profiler BP-203RPEIII (Omron, Kyoto, Japan). The examination room was maintained at 26 ℃. Trained technicians placed 4 pressure cuffs on the participants (one on the upper part of each arm and one on each ankle). Participants were then examined after 10 minutes of rest in the supine position. The device simultaneously recorded bilateral systolic and diastolic blood pressure, ABI, and baPWV, the latter of which was calculated as the ratio of traveled distance (which was automatically estimated from body height) divided by the transit time of the pulse wave between the brachial and posterior tibial arteries. The average of 2-sided baPWV values and 2-sided ABI values were recorded for analysis.

Blood examination

Blood examination (routine blood tests and blood biochemical index tests) were measured using fasting venous blood samples. Routine blood tests were performed using the XN9000 system (Sysmex, Kobe, Japan). Blood biochemical indices, including liver function, renal function, blood lipid profile, fasting blood glucose, HbA1c, and uric acid, were measured using the COBAS 8000 c701 system (Roche, Basel, Switzerland).

Urine examination

Urinary elements were measured using the UF-1000i fully automatic urine analyzer (Sysmex, Kobe, Japan). Urinary chemistry elements were measured using the Siemens Atlas Urine Chemistry Analyzer (Siemens, Erlangen, Germany). The estimated glomerular filtration rate was calculated according to the Cockcroft-Gault formula.

Framingham risk score

The probabilities of CHD and CIS for 10 years are calculated using the following formula:

© Annals of Translational Medicine. All rights reserved.
https://dx.doi.org/10.21037/atm-22-1916

$$P = 1 - Sb(10) \exp(L\text{-}M) \quad [1]$$

where Sb is the probability of survival free for 10 years, L is the linear combination of an individual's risk factor values and the corresponding coefficients, and M is simply L evaluated at the mean of all the covariates.

For female probability of CHD, M = 9.92545, S(10) = 0.96246, and L are calculated as follows:

L = 0.33766 × Age – 0.00268 × Age2 – 0.26138 (if cholesterol < 160) + 0.20771 (if 200 ≤ cholesterol < 240) + 0.24385 (if 240 ≤ cholesterol < 280) + 0.53513 (if cholesterol ≥ 280) + 0.84312 (if HDL-C <35) + 0.37796 (if 35 ≤ HDL-C < 45) + 0.19785 (if 45 ≤ HDL-C < 50) - 0.42951 (if HDL-C ≥ 60) - 0.53363 (systolic blood pressure [SBP] < 120 and DBP <80) - 0.06773 (130 ≤ SBP < 140 or 85 ≤ diastolic blood pressure [DBP] <90) + 0.26288 (140 ≤ SBP < 160 or 90 ≤ DBP <100) + 0.46573 (SBP ≥ 160 or DBP ≥ 100) + 0.59626 (if diabetic) + 0.29246 (if smoker) [2]

For male probability of CHD, M = 3.0975, S (10) = 0.90015, and L were calculated as follows:

L = 0.04826 × Age – 0.65945 (if cholesterol < 160) + 0.17692 (if 200 ≤ cholesterol < 240) + 0.50539 (if 240 ≤ cholesterol < 280) + 0.65713 (if cholesterol ≥ 280) + 0.49744 (if HDL-C < 35) + 0.24310 (if 35 ≤ HDL-C < 45) - 0.05107 (if 50 ≤ HDL-C < 60) - 0.48660 (if HDL-C ≥ 60) -0.00226 (SBP < 120 and DBP <80) + 0.28320 (130 ≤ SBP < 140 or 85 ≤ DBP <90) + 0.52168 (140 ≤ SBP < 160 or 90 ≤ DBP <100) + 0.61859 (SBP ≥ 160 or DBP ≥ 100) + 0.42839 (if diabetic) + 0.52337 (if smoker) [3]

For female probability of CIS, M = 6.6170719, S (10) = 0.95911, and L were calculated as follows:

L = 0.87938 × (Age/10) + 0.51127 (if smoker) – 0.03035 (if CVD) + 1.20720 (if atrial fibrillation) + 0.39796 (if Age ≥ 65) + 1.07111 (if Age < 65 and diabetic) + 0.06565 (if age ≥ 65 and diabetic) + 0.13085 (if taking antihypertensive drugs) + 0.17234 × (SBP-120)/10 (if taking antihypertensive drugs) + 0.11303 × (SBP-120)/10 (if not taking antihypertensive drugs) [4]

For male probability of CIS, M = 4.4227101, S (10) = 0.94451, and L were calculated as follows:

L = 0.49716 × (Age/10) + 0.47254 (if smoker) + 0.45341 (if CVD) + 0.08064 (if atrial fibrillation) + 0.45426 (if age ≥ 65) + 1.35304 (if age < 65 and diabetic) + 0.34385 (if age ≥ 65 and diabetic) + 0.82598 (if taking antihypertensive drugs) + 0.09793 × (SBP-120)/10 (if taking antihypertensive drugs) + 0.27323 × (SBP-120)/10 (if not taking antihypertensive drugs) [5]

© Annals of Translational Medicine. All rights reserved.
https://dx.doi.org/10.21037/atm-22-1916

Table S1 The characteristics of the study population

Variables	Population (N=8,624)
Basic characteristics	
Age, year	51.9±11.4
Sex	
Male	6,152 (71.3)
Female	2,472 (28.7)
Smoker	2,777 (32.2)
Drinker	3,823 (44.3)
BMI, kg/m^2	24.6±3.1
Medical histories	
Atrial fibrillation	19 (0.2)
Chronic obstructive pulmonary disease	11 (0.1)
Coronary heart disease	339 (3.9)
Diabetes	538 (6.2)
Epilepsy	0 (0)
Hemorrhagic stroke	11 (0.1)
Hyperlipidemia	464 (5.4)
Hypertension	1,928 (22.4)
Hypoglycemia	0 (0)
Hypotension	1 (0.01)
Ischemic stroke	339 (3.9)
Myocardiopathy	2 (0.02)
Sick sinus syndrome	2 (0.02)
Transient ischemic attack	1 (0.01)
Valvular heart disease	2 (0.02)
Vascular stiffness	2,427 (28.1)
Medication histories	
Alpha-glucosidase inhibitor	137 (1.6)
Angiotensin converting enzyme inhibitor	144 (1.7)
Angiotensin receptor blocker	652 (7.6)
Antiadrenergic	346 (4)
Antiarrhythmia	6 (0.07)
Anticoagulant	6 (0.07)
Antidiabetic drug	153 (1.8)
Antihypertensive drug	1,431 (16.6)
Antiplatelet	373 (4.3)
Biguanide	143 (1.7)
Calcium channel blocker	820 (9.5)
Dipeptidyl peptidase-4 inhibitor	28 (0.3)
Diuretic	126 (1.5)
Glinide	44 (0.5)
Glucagon-like peptide-1 receptor agonist	0 (0)
Hyperthyroidism	16 (0.2)
Hypothyroidism	110 (1.3)
Insulin	84 (1)
Lipid lowering drug	434 (5)
Nitrates	15 (0.2)
Others antidiabetic drug	77 (0.9)
Others antihypertensive drug	313 (3.6)
Sulfonylurea	74 (0.9)
Thiazolidinedione	20 (0.2)
Uric acid lowering drug	59 (0.7)
Clinical measurements	
Ankle brachial index	1.1±0.07
Brachial-ankle pulse wave velocity, cm/s	1,409.7±256.5
Diastolic blood pressure, mmHg	81.7±11.4
Systolic blood pressure, mmHg	126.3±16.9
Heart rate, time/min	74.5±10.2
Heart rhythm	31 (0.4)
Blood test	
Absolute basophil count, 10^9/L	0.03±0.02
Absolute eosinophil count, 10^9/L	0.2±0.1
Absolute lymphocyte count, 10^9/L	1.9±0.6
Absolute monocyte count, 10^9/L	0.4±0.1
Absolute neutrophil count, 10^9/L	3.3±1.1
Albumin, g/L	44.9±2.6
Alkaline phosphatase, U/L	64±24.9
Alpha-fetoprotein, ng/ml	3.6±1.9
Mean corpuscular hemoglobin concentration, g/L	339.4±12.1
Mean hemoglobin content, pg	30.7±2
Mean RBC volume, fL	90.5±5
Carcinoembryonic antigen, ng/mL	2.1±3.8
Creatinine, μmol/L	75.8±20.6
Direct bilirubin, μmol/L	3.5±1.8
Erythrocyte sedimentation rate, nm/H	5.3±5.8
Estimated glomerular filtration rate, mL/min/1.73 m^2	71.2±19.5
Fasting blood glucose, mmol/L	5.2±1.2
Gamma glutamyl transpeptidase, U/L	37.6±50.2
Globulin, g/L	29.4±3.3
Glutamate-pyruvate transaminase, U/L	23.5±26.2
Glutamic oxaloacetic transaminase, U/L	21.9±13.4
Hematocrit, %	43.5±4
Hemoglobin, g/L	147.6±14.7
Hemoglobin A1C, %	5.8±0.7
High-density lipoprotein cholesterol, mmol/L	1.2±0.3
Low density lipoprotein cholesterol, mmol/L	2.9±0.8
Mean platelet volume, fL	10.1±1.7
Percentage of basophil, %	0.5±0.3
Percentage of eosinophils, %	2.6±2
Percentage of lymphocyte, %	33.8±7.4
Percentage of monocyte, %	7±1.8
Percentage of neutrophils, %	56.2±7.9
Platelet count, 10^9/L	216.6±54.1
Platelet large cell ratio, %	32.3±6.9
Platelet crit, %	0.2±0.06
Platelet distribution width, fL	14.2±2.3
Prostate-specific antigen, ng/mL	1.2±2.1
Red blood cell, 10^{12}/L	4.8±0.5
Total bilirubin, μmol/L	13.0±5.1
Total cholesterol, mmol/L	4.7±0.9
Total protein, g/L	74.4±4.1
Triglyceride, mmol/L	1.6±1.4
Urea, mmol/L	5.2±1.3
Uric acid, mmol/L	352.7±90.1
White blood cell count, 10^9/L	5.9±1.6
Urine test	
Fungi	9 (0.1)
Ketone	111 (1.3)
Nitrite	67 (0.8)
Pathological cast count, /mL	0.1±0.2
Small round epithelial cell	177 (2.1)
Urinary cast count, /mL	0.3±0.4
Urinary erythrocyte morphology	479 (5.6)
Urinary occult blood	1,106 (12.8)
Urine bilirubin	293 (3.4)
Urine crystal	676 (7.8)
Urine epithelial cells count, /mL	8.6±18.6
Urine glucose	90 (1)
Urine pH	6.2±0.7
Urine protein	351 (4.1)
Urine red blood cell count, /mL	20.9±374.9
Urine specific gravity	1±0.007
Urine white blood cell	597 (6.9)
Urine white blood cell count, /mL	21.8±179
Urobilinogen	8 (0.09)

Continuous variables were described as mean ± SD. Categorical variables are described as number (percentage). BMI, body mass index; RBC, red blood cell; SD, standard deviation.

© Annals of Translational Medicine. All rights reserved.
https://dx.doi.org/10.21037/atm-22-1916

Table S2 Characteristics of the CHD dataset (N=7,698)

Variables	Training (N=5,136)	Random testing (N=1,284)	Sequential testing (N=1,278)
Basic characteristics			
Age	52.8±9.8	52.5±10.1	53.2±10.5
Male	3,721 (72.4)	900 (70.1)	927 (72.5)
Smoker	1,686 (32.8)	395 (30.8)	466 (36.5)
Drinker	2,245 (43.7)	538 (41.9)	670 (52.4)
BMI, kg/m^2	24.7±3.03	24.6±2.98	24.6±3.1
Cardiovascular medical history			
CHD	219 (4.3)	57 (4.4)	63 (4.9)
CIS	25 (0.5)	5 (0.4)	7 (0.5)
Clinical measurements			
SBP, mmHg	126.7±16.6	126.5±16.7	126.7±17.8
DBP, mmHg	82.4±11.5	81.9±11.3	81±11.4
LDL-C, mmol/L	2.9±0.8	2.9±0.8	2.8±0.8
HDL-C, mmol/L	1.2±0.3	1.2±0.3	1.2±0.3
TC, mmol/L	4.8±0.9	4.8±0.9	4.5±0.9
TG, mmol/L	1.7±1.4	1.7±1.5	1.7±1.6
FBG, mmol/L	5.3±1.2	5.3±1.2	5.2±1.2
HbA1c, %	5.8±0.7	5.8±0.7	5.7±0.7
baPWV, cm/s	1,416.6±247.5	1,411.2±241.1	1,430.7±270.7
<1,400	2,807 (54.7)	711 (55.4)	664 (52)
1,400–1,800	1,984 (38.6)	476 (37.1)	507 (39.7)
≥1,800	345 (6.7)	97 (7.6)	107 (8.4)
ABI	1.1±0.07	1.1±0.07	1.1±0.07
<0.9	28 (0.5)	15 (1.2)	7 (0.5)
0.9–1.4	5,100 (99.3)	1,267 (98.7)	1,270 (99.4)
≥1.4	8 (0.2)	2 (0.2)	1 (0.1)

Continuous variables are described as mean ± SD. Categorical variables are described as number (percentage). CHD, coronary heart disease; BMI, body mass index; CIS, cerebral ischemic stroke; SBP, systolic blood pressure; DBP, diastolic blood pressure; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; TC, total cholesterol; TG, triglyceride; FBG, fasting blood glucose; HbA1c, hemoglobin A1C; baPWV, brachial-ankle pulse wave velocity; ABI, ankle brachial index; SD, standard deviation.

© Annals of Translational Medicine. All rights reserved.
https://dx.doi.org/10.21037/atm-22-1916

Table S3 Characteristics of the CIS dataset (N=8,322)

Variables	Training (N=5,503)	Random testing (N=1,375)	Sequential testing (N=1,444)
Basic characteristics			
Age	51.5±10.9	51.6±11.2	50.7±12.1
Male	3,899 (70.9)	961 (69.9)	1,027 (71.1)
Smoker	1,752 (31.8)	418 (30.4)	491 (34)
Drinker	2,416 (43.9)	571 (41.5)	740 (51.2)
BMI, kg/m^2	24.6±3.1	24.5±3.1	24.4±3.2
Cardiovascular medical history			
CHD	26 (0.5)	4 (0.3)	7 (0.5)
CIS	230 (4.2)	52 (3.8)	57 (3.9)
Clinical measurements			
SBP, mmHg	126.2±16.5	126.2±17.2	125.6±17.6
DBP, mmHg	82±11.4	82±11.6	80.4±11.3
LDL-C, mmol/L	2.9±0.8	2.9±0.7	2.8±0.8
HDL-C, mmol/L	1.2±0.3	1.2±0.3	1.2±0.3
TC, mmol/L	4.8±0.9	4.7±0.9	4.5±0.9
TG, mmol/L	1.6±1.4	1.7±1.5	1.6±1.5
FBG, mmol/L	5.2±1.2	5.2±1.2	5.1±1.1
HbA1c, %	5.8±0.7	5.8±0.7	5.7±0.7
baPWV, cm/s	1401.8±247.7	1408.7±252.7	1400.7±272.8
<1,400	3,124 (56.8)	780 (56.7)	823 (57)
1,400–1,800	2,017 (36.7)	494 (35.9)	514 (35.6)
>1,800	362 (6.6)	101 (7.3)	107 (7.4)
ABI	1.1±0.07	1.1±0.06	1.1±0.07
<0.9	42 (0.8)	6 (0.4)	7 (0.5)
0.9–1.4	5,453 (99.1)	1,368 (99.5)	1,436 (99.4)
>1.4	8 (0.1)	1 (0.1)	1 (0.1)

Continuous variables are described as mean ± SD. Categorical variables are described as number (percentage). CIS, cerebral ischemic stroke; BMI, body mass index; CHD, coronary heart disease; SBP, systolic blood pressure; DBP, diastolic blood pressure; LDL-C, low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein cholesterol; TC, total cholesterol; TG, triglyceride; FBG, fasting blood glucose; HbA1c, hemoglobin A1C; baPWV, brachial-ankle pulse wave velocity; ABI, ankle brachial index; SD, standard deviation.

© Annals of Translational Medicine. All rights reserved.
https://dx.doi.org/10.21037/atm-22-1916

Table S4 Selected features by ANOVA, RFE, and Boruta

Disease	Overlap features	Unique features
CHD	TC, baPWV, AMC, RBC count, LDL-C, PSA, creatinine, monocyte, SBP, MCV, CEA, eGFR, platelet count, FBG, age, urea	UA, HDL-C, lymphocyte, HbA1c, TG, heart rate, BMI
Ischemic stroke		MCH, WBC count, ABI, ANC, ALP, globulin

Features were collected into category "overlap features" if they presented both in CHD and ischemic. Otherwise, features will be collected into "unique features". ANOVA, analysis of variance; RFE, recursive feature elimination; CHD, coronary heart disease; TC, total cholesterol; baPWV, brachial-ankle pulse wave velocity; AMC, absolute monocyte count; RBC, red blood cell; LDL-C, low-density lipoprotein cholesterol; PSA, prostate specific antigen; SBP, systolic blood pressure; MCV, average red blood cell volume; CEA, carcinoembryonic antigen; eGFR, estimated glomerular filtration rate; FBG, fasting blood glucose; UA, uric acid; HDL-C, high-density lipoprotein cholesterol; HbA1c, hemoglobin A1C; TG, triglyceride; BMI, body mass index; MCH, average hemoglobin content; WBC, white blood cell; ABI, ankle brachial index; ANC, absolute neutrophil count; ALP, alkaline phosphatase.

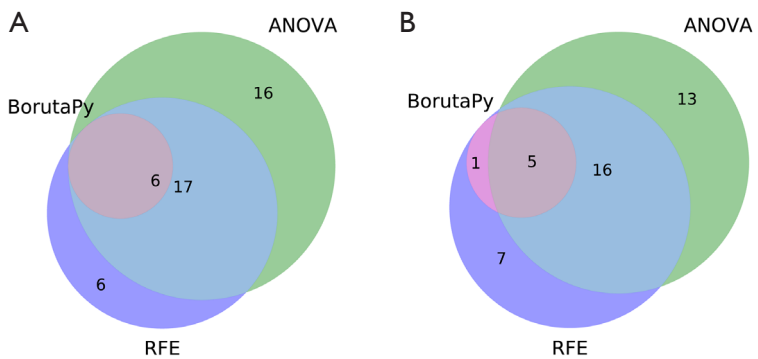

Figure S1 The results of three feature selection methods. (A) Coronary heart disease dataset. (B) Ischemic stroke dataset. ANOVA, analysis of variance; RFE, recursive feature elimination; BorutaPy, boruta for Python.

© Annals of Translational Medicine. All rights reserved.
https://dx.doi.org/10.21037/atm-22-1916

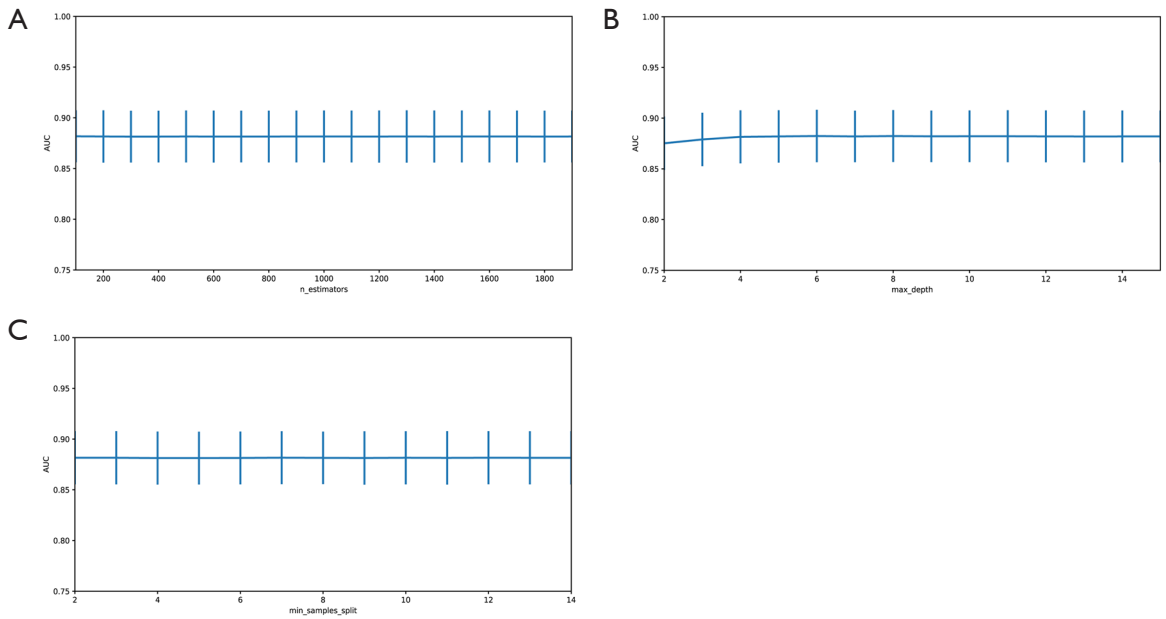

Figure S2 GridSearchCV analysis results of coronary heart disease classification. (A) Model performance with different estimators. (B) Model performance with different maximum depths. (C) Model performance with different minimum sample split. AUC, area under the receiver operating characteristic curve.

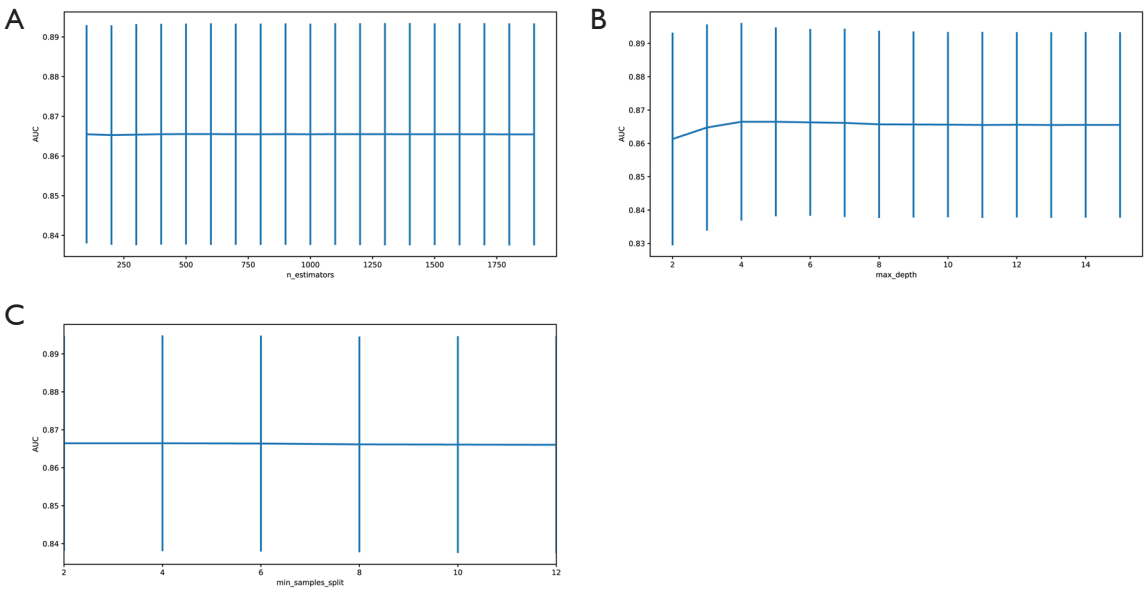

Figure S3 GridSearchCV analysis results of ischemic stroke classification. (A) Model performance with different estimators. (B) Model performance with different maximum depths. (C) Model performance with different minimum sample split. AUC, area under the receiver operating characteristic curve.

© Annals of Translational Medicine. All rights reserved.
https://dx.doi.org/10.21037/atm-22-1916

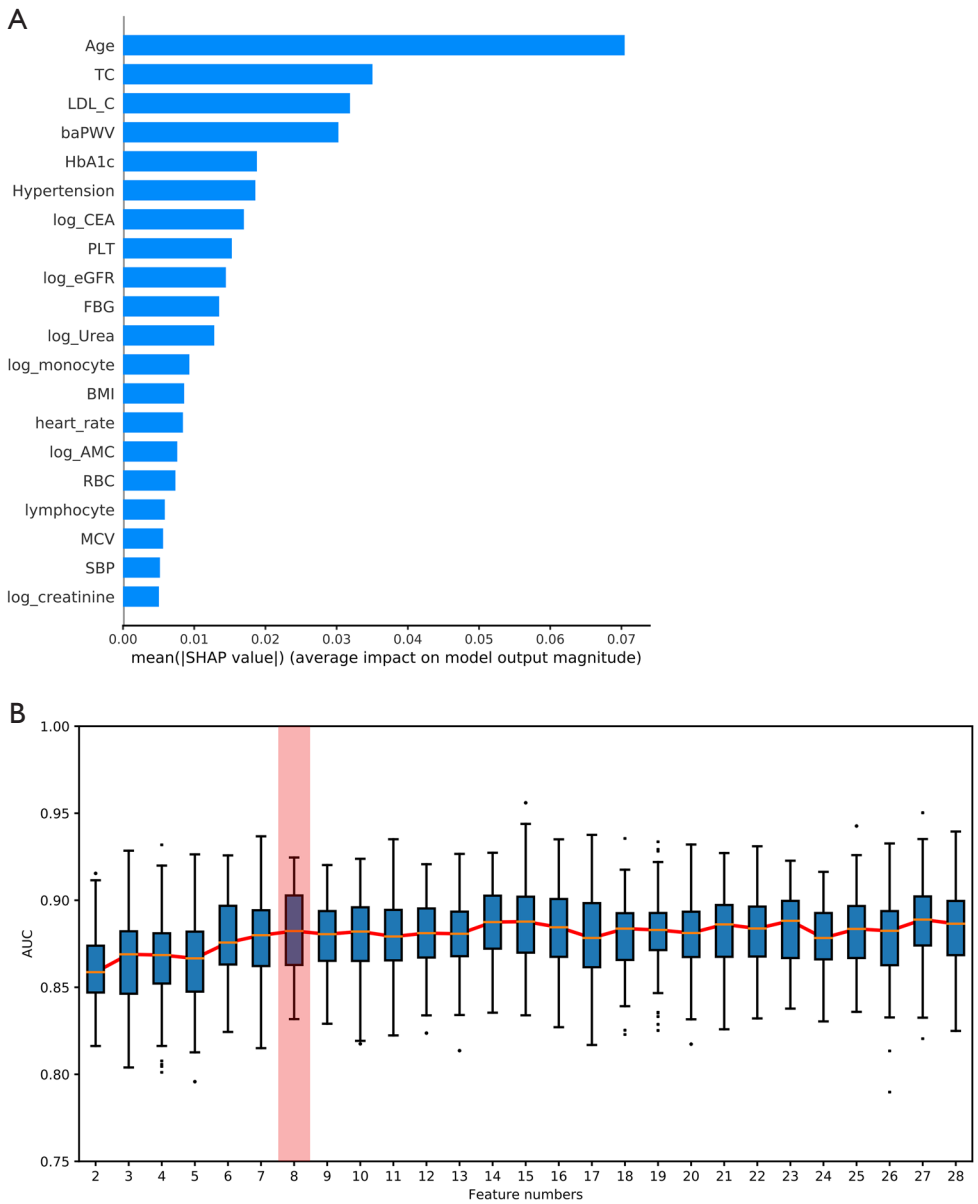

Figure S4 Optimization of the CHD model. (A) Mean absolute SHAP values of each feature in the CHD model during the optimization process. (B) The minimum features required in the CHD model. The inflection point is marked by the red shadow. TC, total cholesterol; LDL-C, low-density lipoprotein cholesterol; baPWV, brachial-ankle pulse wave velocity; HbA1c, hemoglobin A1C; CEA, carcinoembryonic antigen; PLT, platelet; eGFR, estimated glomerular filtration rate; FBG, fasting blood glucose; BMI, body mass index; AMC, absolute monocyte count; RBC, red blood cell; MCV, average red blood cell volume; SBP, systolic blood pressure; SHAP, shapley additive explanations; AUC, area under the receiver operating characteristic curve; CHD, coronary heart disease.

© Annals of Translational Medicine. All rights reserved.
https://dx.doi.org/10.21037/atm-22-1916

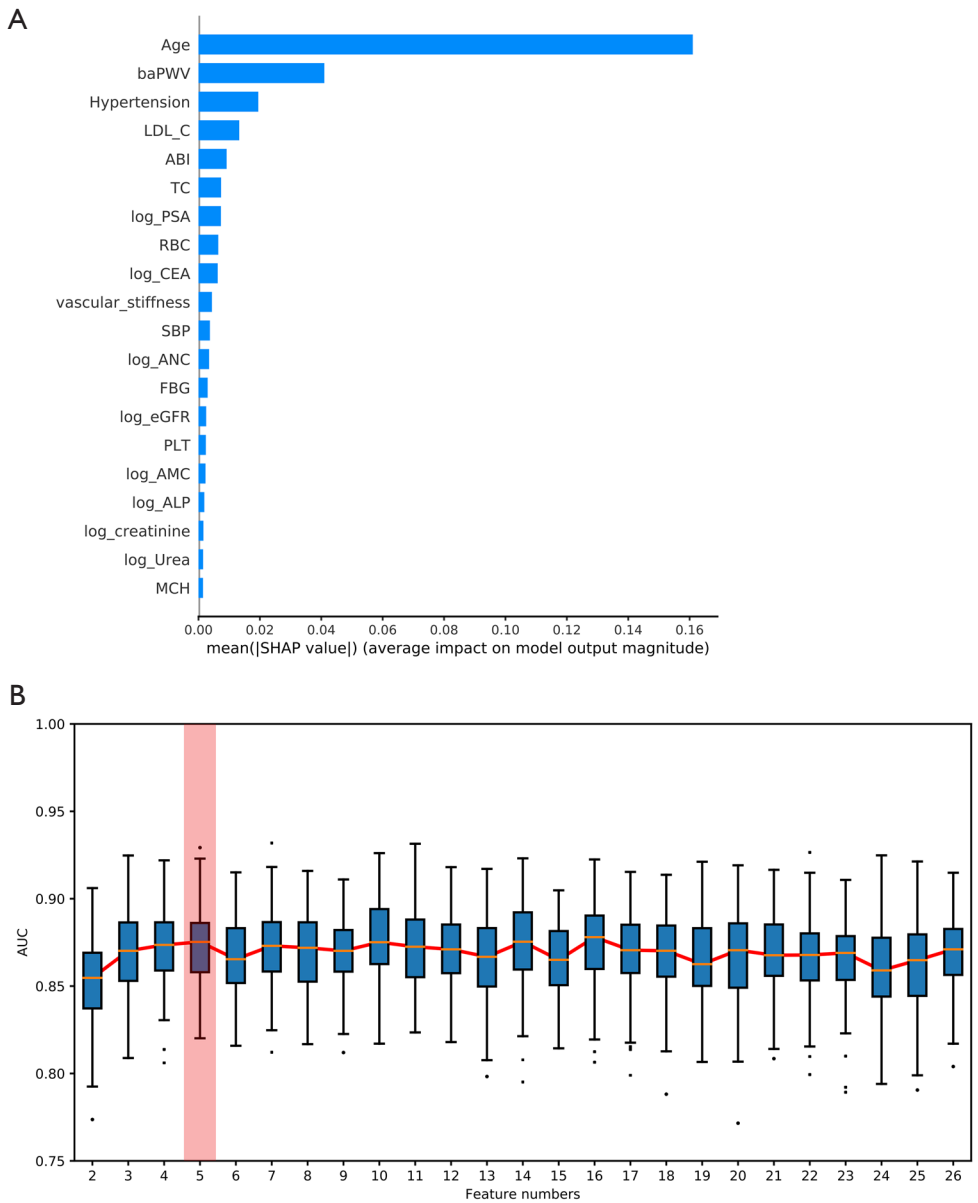

Figure S5 Optimization of the CIS model. (A) Mean absolute SHAP values of each feature in the CIS model during the optimization process. (B) The minimum features required in the CIS model. The inflection point is marked by the red shadow. baPWV, brachial-ankle pulse wave velocity; LDL-C, low-density lipoprotein cholesterol; ABI, ankle brachial index; TC, total cholesterol; PSA, prostate specific antigen; RBC, red blood cell; CEA, carcinoembryonic antigen; SBP, systolic blood pressure; ANC, absolute neutrophil count; FBG, fasting blood glucose; eGFR, estimated glomerular filtration rate; PLT, platelet; AMC, absolute monocyte count; ALP, alkaline phosphatase; MCH, average hemoglobin content; SHAP, shapley additive explanations; AUC, area under the receiver operating characteristic curve; CIS, cerebral ischemic stroke.

© Annals of Translational Medicine. All rights reserved.
https://dx.doi.org/10.21037/atm-22-1916

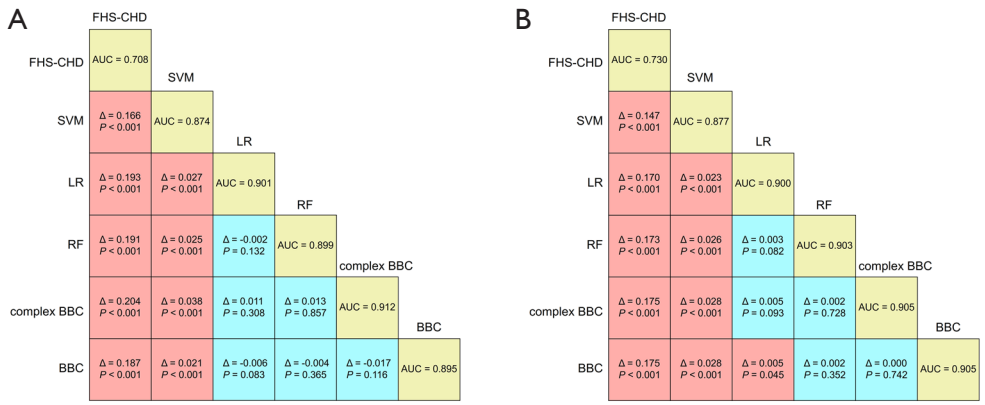

Figure S6 Statistical analysis of the AUCs of the CHD models. (A) Performance of the models using random test data. (B) Performance of the models using sequential test data. Δ represents the difference in AUCs; P represents the result of the ROC curve difference test; the yellow squares represent the AUC values; the square would be red if P < 0.05, and the square would be blue if P≥0.05. FHS-CHD, Framingham risk score of coronary heart disease; SVM, support vector machine; LR, logistic regression; RF, random forest; BBC, balanced bagging classifier; AUC, the area under the receiver operating characteristic curve; CHD, coronary heart disease.

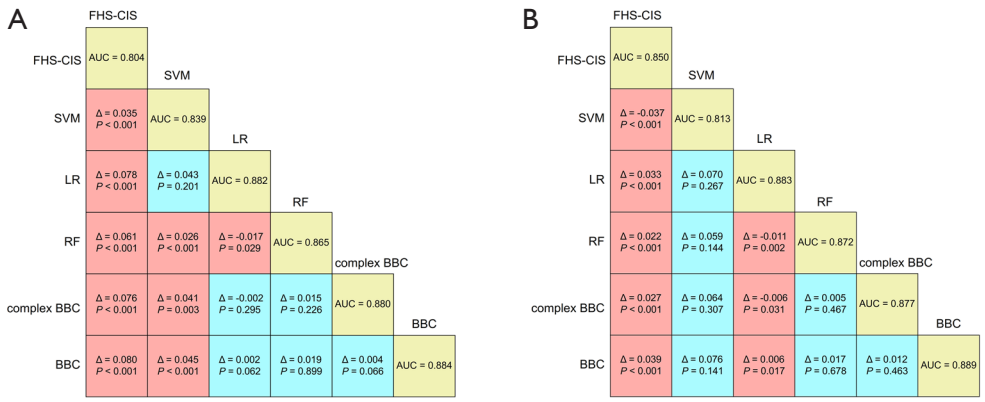

Figure S7 Statistical analysis of the AUCs of the CIS models. (A) Performance of the models using random test data. (B) Performance of the models using sequential test data. Δ represents the difference in AUCs; P represents the result of the ROC curve difference test; the yellow squares represent the AUC values; the square would be red if P < 0.05, and the square would be blue if P≥0.05. FHS-CIS, Framingham risk score of cerebral ischemic stroke; SVM, support vector machine; LR, logistic regression; RF, random forest; BBC, balanced bagging classifier; AUC, the area under the receiver operating characteristic curve; CIS, cerebral ischemic stroke.

© Annals of Translational Medicine. All rights reserved.
https://dx.doi.org/10.21037/atm-22-1916

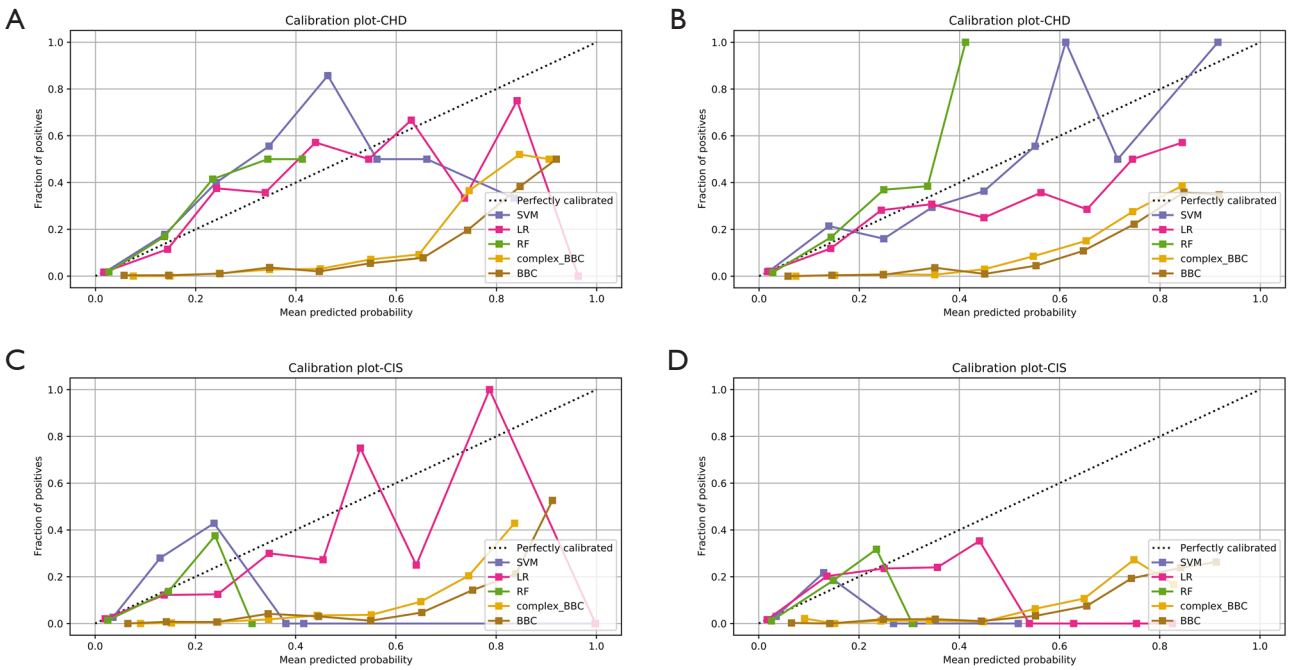

Figure S8 Calibration of each model using different datasets. (A) Random test data of CHD. (B) Sequential test data of CHD. (C) Random test data of CIS. (D) Sequential test data of CIS. CHD, coronary heart disease; SVM, support vector machine; LR, logistic regression; RF, random forest; BBC, balanced bagging classifier; CIS, cerebral ischemic stroke.

© Annals of Translational Medicine. All rights reserved.
https://dx.doi.org/10.21037/atm-22-1916

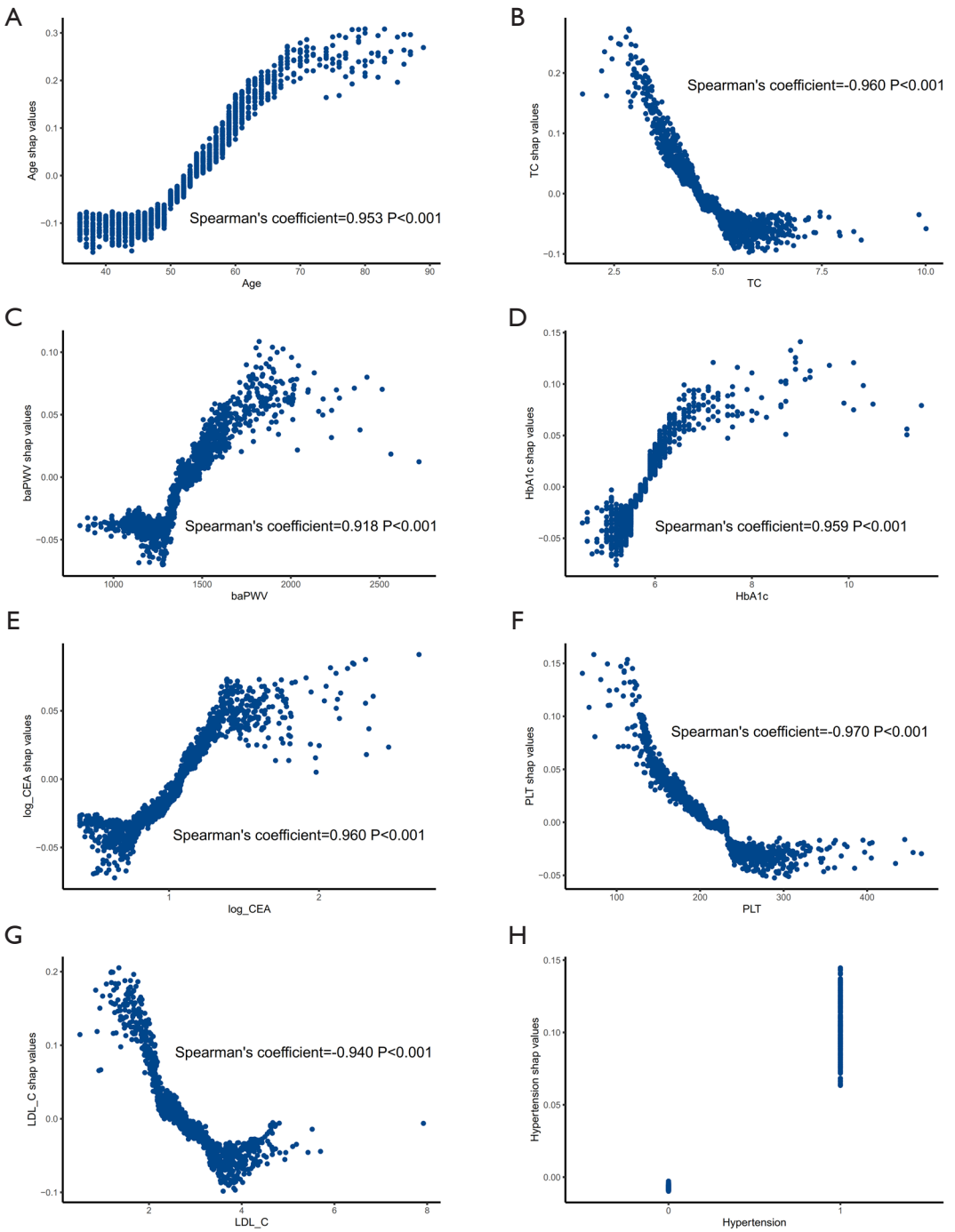

Figure S9 SHAP dependence plot of the coronary heart disease model. The x-axis represents the value of the feature. The y-axis shows the corresponding SHAP value, and represents the feature contribution on model output. (A-H) show the age, TC, baPWV, HbA1c, CEA, platelet count, LDL-C, and hypertension, respectively. TC, total cholesterol; baPWV, brachial-ankle pulse wave velocity; HbA1c, hemoglobin A1C; CEA, carcinoembryonic antigen; PLT, platelet; LDL-C, low-density lipoprotein cholesterol.

© Annals of Translational Medicine. All rights reserved.
https://dx.doi.org/10.21037/atm-22-1916

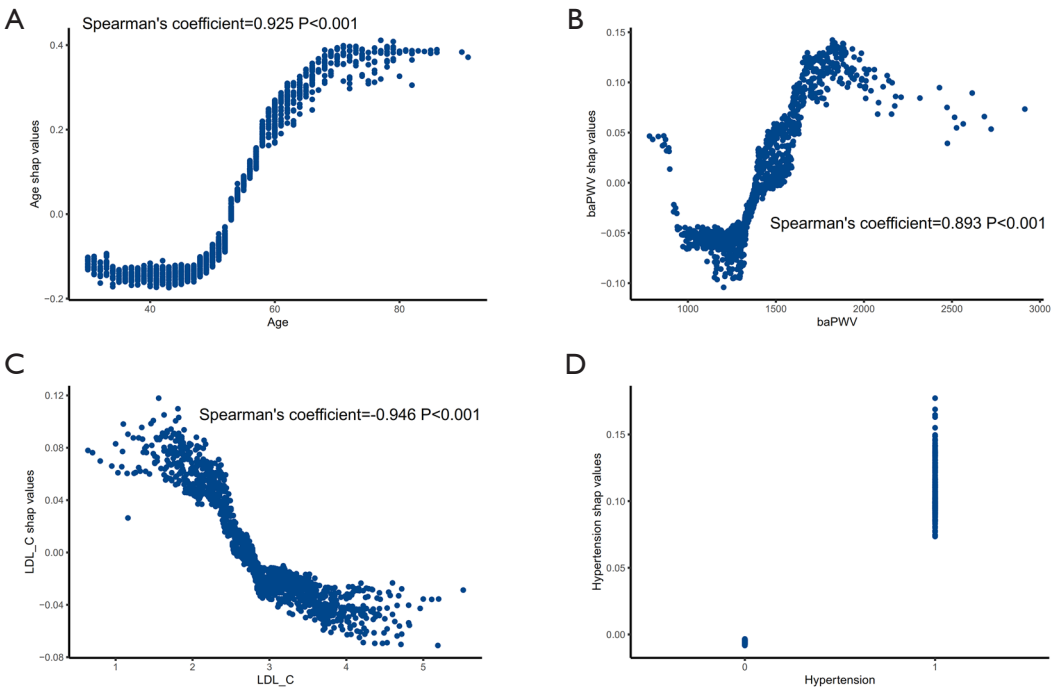

Figure S10 SHAP dependence plot of the ischemic stroke model. The x-axis represents the value of the feature. The y-axis shows the corresponding SHAP value, which represents the feature contribution on model output. (A-D) show the age, baPWV, LDL-C, and hypertension, respectively. baPWV, brachial-ankle pulse wave velocity; LDL-C, low-density lipoprotein cholesterol; SHAP, Shapley Additive exPlanations.

© Annals of Translational Medicine. All rights reserved.
https://dx.doi.org/10.21037/atm-22-1916

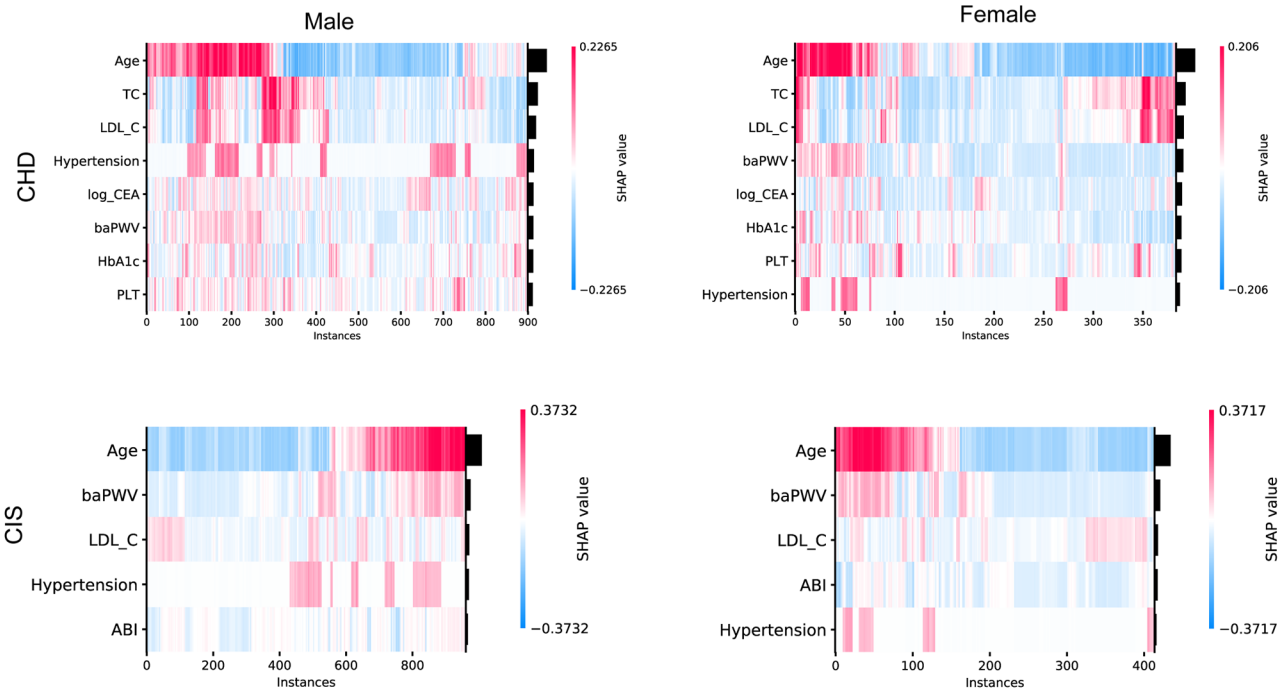

Figure S11 Analysis of gender differenced for features used in the models. CHD, coronary heart disease; CIS, cerebral ischemic stroke; TC, total cholesterol; LDL-C, low density lipoprotein cholesterol; CEA, carcinoembryonic antigen; baPWV, brachial-ankle pulse wave velocity; HbA1c, hemoglobin A1c; PLT, platelet; ABI, ankle-brachial index.

© Annals of Translational Medicine. All rights reserved.
https://dx.doi.org/10.21037/atm-22-1916

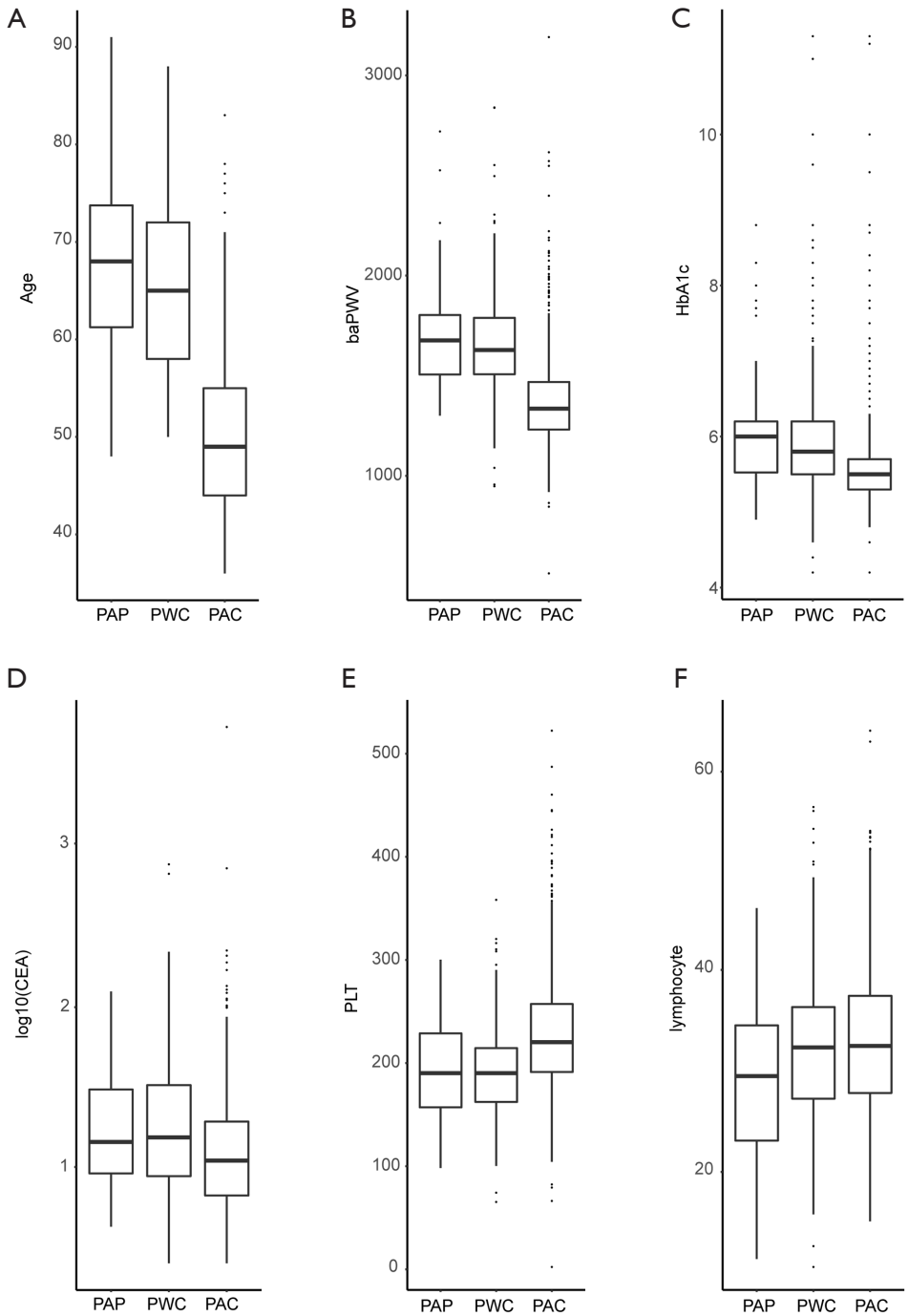

Figure S12 The dissimilarity characteristics between predicted accurate samples and predicted wrong controls in the coronary heart disease model. The y-axis represents characteristics, and (A-F) show the age, baPWV, HbA1c, CEA, platelet count, and percentage of lymphocytes, respectively. PAP, predicted accurate patients; PWC, predicted wrong controls; PAC, predicted accurate controls; baPWV, brachial-ankle pulse wave velocity; HbA1c, hemoglobin A1C; CEA, carcinoembryonic antigen; PLT, platelet.

© Annals of Translational Medicine. All rights reserved.
https://dx.doi.org/10.21037/atm-22-1916

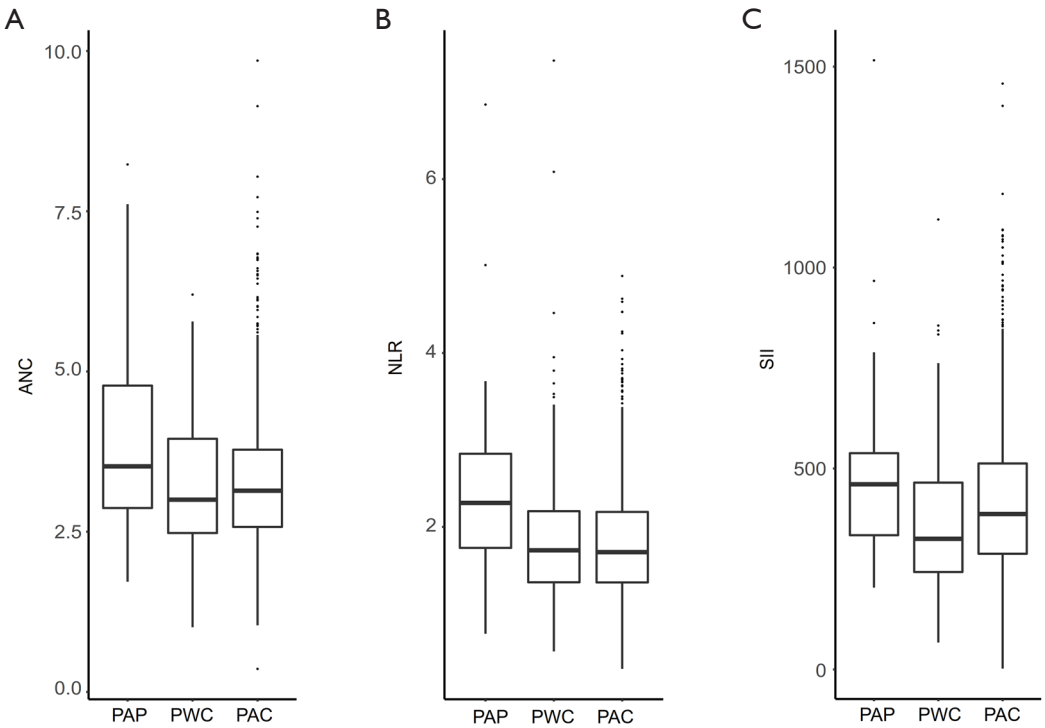

Figure S13 The value of inflammatory markers between predicted accurate samples and predicted wrong controls in the CHD model. The y-axis represents the characteristics, and (A-C) show the ANC, NLR, and SII, respectively. ANC, absolute neutrophil count; PAP, predicted accurate patients; PWC, predicted wrong controls; PAC, predicted accurate controls; NLR, neutrophil/lymphocyte ratio; SII, systemic immune inflammation index.

© Annals of Translational Medicine. All rights reserved.
https://dx.doi.org/10.21037/atm-22-1916