

# Pulse Rate Deficit Identifies Distinct Persistent Atrial Fibrillation Phenotypes via Unsupervised Clustering: A Single-Center Prospective Study

xiang xue

The First Affiliated Hospital of Nanjing Medical University

Kai Li

Donghai County People's Hospital

Cheng Cai

The First Affiliated Hospital of Nanjing Medical University

Hailei Liu

The First Affiliated Hospital of Nanjing Medical University

Minglong Chen

[chenminglong@njmu.edu.cn](mailto:chenminglong@njmu.edu.cn)

The First Affiliated Hospital of Nanjing Medical University

---

## Research Article

**Keywords:** Atrial fibrillation, Pulse rate deficit, Phenotyping, Unsupervised clustering, Catheter ablation

**Posted Date:** October 5th, 2026

**DOI:** <https://doi.org/10.21203/rs.3.rs-10826079/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

**Additional Declarations:** No competing interests reported.

---

1 Pulse Rate Deficit Identifies Distinct Persistent Atrial Fibrillation Phenotypes via Unsupervised  
2 Clustering: A Single-Center Prospective Study

3

4 Xiang Xue<sup>1,2</sup>, Kai Li<sup>3</sup>, Cheng Cai<sup>1</sup>, Hailei Liu<sup>1\*</sup>, Minglong Chen<sup>1\*</sup>

5

6 1 Department of Cardiology, The First Affiliated Hospital of Nanjing Medical University, Nanjing,  
7 Jiangsu, China

8 2 Department of Cardiology, Changzhou Geriatric Hospital Affiliated to Soochow University  
9 (Changzhou No. 7 People's Hospital), Changzhou, Jiangsu, China

10 3 Department of Cardiology, Donghai County People's Hospital, Lianyungang, Jiangsu, China

11

12 \*Corresponding authors:

13 Minglong Chen, Department of Cardiology, The First Affiliated Hospital of Nanjing Medical  
14 University, Nanjing, Jiangsu, China. Email: chenminglong@njmu.edu.cn

15 Hailei Liu, Department of Cardiology, The First Affiliated Hospital of Nanjing Medical University,  
16 Nanjing, Jiangsu, China. Email: liuhailei@njmu.edu.cn

17

18 Abstract

19

20 Background: Persistent atrial fibrillation (AF) is phenotypically heterogeneous. Existing  
21 phenotyping approaches rely predominantly on costly imaging or biomarker panels, limiting  
22 routine clinical applicability. Pulse rate deficit (PRD), an easily measurable clinical sign, is  
23 associated with exercise intolerance and impaired hemodynamics. Whether PRD can classify AF  
24 phenotypes and predict the response to catheter ablation remains unknown.

25

26   Methods: Consecutive patients with persistent AF undergoing first catheter ablation were  
27   prospectively enrolled at a single center; 103 formed the final clustering cohort. PRD was  
28   measured as |heart rate - pulse rate| (bpm) by a single investigator using sequential cardiac  
29   auscultation (60 s) and radial artery palpation (60 s), with measurement order randomized per  
30   patient. Pearson correlation analysis assessed the clinical correlates of PRD. Unsupervised  
31   K-means clustering ( $K=5$ ) was applied to 7 standardized features (PRD, heart rate, pulse rate, LVEF,  
32   NT-proBNP, eGFR, age; random\_state=42). Clustering stability was evaluated across 10 random  
33   seeds using the adjusted Rand index (*ARI*) and reclassification rate.

34

35   Results: Of 108 enrolled patients, 103 (mean age  $63.2 \pm 9.8$  years, 22.3% female) formed the final  
36   clustering cohort. PRD correlated significantly with heart rate ( $r=+0.570$ ,  $P<0.001$ ), LVEF ( $r=-0.382$ ,  
37    $P<0.001$ ), heart failure ( $r=+0.342$ ,  $P<0.001$ ) and WBC ( $r=+0.400$ ,  $P<0.001$ ). Five clusters were  
38   identified: C0 (markedly elevated PRD, preserved LVEF,  $n=16$ ), C1 (low PRD, young,  $n=45$ ), C2  
39   (extreme phenotype: youngest, highest PRD, lowest LVEF, 100% heart failure,  $n=4$ ), C3 (oldest,  
40   low PRD,  $n=29$ ) and C4 (highest NT-proBNP,  $n=9$ ). The silhouette score at  $K=5$  was 0.196. C2  
41   remained stable in 9/10 seeds (mean *ARI*=0.724, mean reclassification rate=11.4%). In  
42   exploratory analyses, recurrence rates did not differ significantly across clusters ( $P=0.386$ ).

43

44   Conclusions: PRD is a valid, bedside-available phenotypic marker that identifies distinct AF  
45   subgroups, including a reproducible extreme subtype (C2: young, high PRD, reduced LVEF, 100%  
46   heart failure). Its primary value lies in cross-sectional phenotyping rather than prognosis

prediction. Larger studies are needed to validate these phenotypes and their prognostic implications.

Keywords: Atrial fibrillation; Pulse rate deficit; Phenotyping; Unsupervised clustering; Catheter ablation

## Background

Atrial fibrillation (AF) is a heterogeneous disease with diverse clinical manifestations and a substantial global health burden [1]. Although catheter ablation is an established treatment [2,3], its success rate varies substantially across patients, especially in patients with persistent AF [2]. Phenotyping AF patients using multidimensional clinical data has emerged as a promising approach to identify subgroups with distinct mechanisms and treatment responses [4,5]. However, existing phenotyping studies have predominantly relied on costly imaging or biomarker panels [4,6], limiting their routine clinical applicability. A simple, bedside-available parameter capable of contributing to AF phenotyping would be of considerable practical value.

Pulse rate deficit (PRD), defined as the absolute difference between heart rate (by auscultation) and pulse rate (by palpation), reflects the number of cardiac cycles with insufficient stroke volume to generate a peripheral pulse, a direct consequence of RR interval irregularity in AF [7].

As a clinical sign, pulse deficit has been recognized for decades as a simple, bedside marker of ineffective cardiac output in AF, which is found to be linked to exercise intolerance and impaired hemodynamics [7]. Impaired hemodynamics have been found to be associated with AF recurrence [8]. While data-driven approaches have been explored to classify AF phenotypes with complex parameters [4,5,9,10], whether PRD, a readily available clinical parameter, can identify clinically meaningful AF subgroups remains unknown.

We hypothesized that integrating PRD with multidimensional clinical parameters could identify novel AF phenotypes via unsupervised machine learning. The primary objectives were to: (1) assess the clinical correlates of PRD by correlating it with established clinical indicators; (2) identify clinically meaningful AF phenotypes using unsupervised clustering; and (3) explore the association between PRD-derived phenotypes and post-ablation recurrence.

## Methods

### Study Population

This single-center prospective study consecutively enrolled patients with persistent AF undergoing first catheter ablation at the First Affiliated Hospital with Nanjing Medical University, between August 2025 and December 2025. Inclusion criteria: (1) age  $\geq 18$  years; (2) persistent AF, defined as AF sustained beyond 7 days or requiring pharmacological/electrical cardioversion for

termination, according to current guidelines [2]; (3) first ablation procedure. Exclusion criteria: (1) hypertrophic cardiomyopathy; (2) permanent pacemaker implantation; (3) hyperthyroidism or hypothyroidism; (4) severe valvular disease (including moderate-to-severe aortic stenosis or regurgitation, and moderate-to-severe mitral stenosis); (5) missing PRD measurement, due to unclear heart sounds or impalpable pulse; (6) incomplete baseline data. Patients with impalpable radial pulse were further evaluated by upper extremity vascular ultrasound to exclude peripheral arterial disease. The study protocol was approved by the Institutional Review Board of the First Affiliated Hospital of Nanjing Medical University, and all patients provided written informed consent.

#### Baseline Characteristics

Baseline characteristics were collected at admission, which included: (1) demographics (age, gender, BMI); (2) AF-related parameters (AF duration, EHRA symptom score); (3) vital signs (heart rate, pulse rate); (4) comorbidities (hypertension, diabetes mellitus, heart failure, prior stroke/TIA/thromboembolism, coronary artery disease or peripheral vascular disease); and (5) medications (anticoagulants, RAAS inhibitors, beta-blockers, statins).

Additional data were collected before the index procedure, which included (1) echocardiographic parameters (LVEF, left ventricular end-diastolic diameter, left ventricular end-systolic diameter, left atrial diameter); (2) laboratory indicators (NT-proBNP, eGFR, WBC, neutrophil count, hemoglobin, albumin, indirect bilirubin);

Additional variables are listed in Table 1. All measurements were performed by blinded investigators.

113

#### 114 PRD Measurement

115

116 PRD was measured by a single investigator (blinded to clinical data) using cardiac auscultation  
117 (60 seconds at the cardiac apex) and radial artery palpation (60 seconds), performed sequentially  
118 with randomized order. The measurement order (auscultation first vs. palpation first) was  
119 determined using a computer-generated random sequence prior to each patient encounter. PRD  
120 was calculated as  $|\text{heart rate} - \text{pulse rate}|$  (bpm). All measurements were performed at the first  
121 medical contact before ablation and during persistent AF (confirmed by 12-lead ECG).

122

#### 123 Ablation procedure

124 Ablation was performed following the STABLE-SR strategy [11,12]. Circumferential pulmonary  
125 vein isolation was performed first. Sinus rhythm was then restored by electrical cardioversion,  
126 and high-density bipolar voltage mapping of the left atrium was performed during sinus rhythm.  
127 Left atrial low voltage area (LVA) was defined as a bipolar electrogram amplitude  $<0.50$  mV at  $\geq 3$   
128 adjacent points during sinus rhythm and were modified when present. Detailed LVA modification  
129 strategy was previously reported [12].

130

131

#### 132 Follow-up and Recurrence Definition

133

134 All patients were followed for 6 months after ablation. Regular 12-lead ECG and 24-hour Holter

monitoring was performed at 6 months. Antiarrhythmic drugs were prescribed for 3 months post-ablation unless contraindicated and discontinued thereafter [2,3]. Anticoagulation was prescribed for at least 3 months and its continuation was left to the discretion of physicians. Recurrence was defined as any documented episode of atrial fibrillation, atrial flutter, or atrial tachycardia lasting  $\geq 30$  seconds on 12-lead ECG or 24-hour Holter monitoring after a 3-month blanking period [3].

## Statistical Analysis

Correlation analysis: Pearson correlation was used to assess PRD's correlations with clinical variables (reported as  $r$  and  $P$  values), among all 104 patients with available PRD data. One additional patient who had PRD data but was excluded from clustering due to incomplete key variables was therefore included in the correlation analysis, while the clustering analysis comprised 103 patients.

Clustering analysis: We applied K-means clustering to identify patient subgroups, building on a growing body of unsupervised phenotyping studies in AF [4,5,9,10] and established methodologic approaches for unsupervised learning in cardiovascular research [13]. Feature selection for clustering was guided by clinical plausibility and pathophysiological considerations, aiming to capture multidimensional aspects of AF without overfitting the modest sample size ( $n=103$ ). Seven continuous variables were chosen a priori to represent four domains: (1) pulse deficit (PRD, the primary variable); (2) rate-related parameters (heart rate, pulse rate); (3) cardiac function and

remodeling (LVEF, NT-proBNP); and (4) comorbidity burden (eGFR, age).

We deliberately excluded other available variables (e.g., LAD, left atrial low voltage, inflammatory markers, categorical comorbidities) for three reasons: (1) binary/categorical variables are suboptimal for Euclidean distance-based K-means clustering; (2) including too many features would increase the risk of overfitting given the sample size; and (3) we conducted two sensitivity analyses — adding left atrial diameter (LAD) and adding left atrial low voltage separately to the 7-feature model (8-feature models) — and compared their silhouette scores and Adjusted Rand Indices against the primary 7-feature clustering to assess whether these variables materially altered cluster stability. The 7-feature model was retained for parsimony and reproducibility on the basis of these analyses (reported in Results).

All continuous variables were standardized using StandardScaler (mean=0, SD=1) to prevent any single variable from dominating the distance calculation. The number of clusters was determined by silhouette analysis [14] across  $K=2-7$ ; while  $K=2$  yielded the highest silhouette score (0.257),  $K=5$  was selected to maximize clinical interpretability (silhouette score=0.196), as binary partitioning would be insufficient to capture the phenotypic diversity of AF. To ensure reproducibility, the random seed was fixed at 42. K-means was initialized with k-means++ ( $n\_init=10$ ,  $max\_iter=300$ ). We acknowledge that selecting  $K=5$  rather than the statistically favored  $K=2$  involved partial subjective judgment, and our findings require external validation.

Clustering stability: Clustering stability was assessed by running 10 different random seeds (1–10)

and calculating: (1) the Adjusted Rand Index (*ARI*) [15] between each partition and the seed-42 partition; and (2) the reclassification rate (proportion of patients assigned to a different cluster relative to seed 42, after optimal label matching using the Hungarian algorithm).

Exploratory prognostic analysis: As a secondary objective, we performed an exploratory analysis of post-ablation recurrence. Cluster-based recurrence rates were compared using the chi-square test. Logistic regression was used to evaluate the association between PRD (continuous variable, standardized) and recurrence, with adjustment for clinically relevant predefined variables (LAD, age, gender, LVEF, NT-proBNP, and eGFR; total 7 variables including PRD). Variables were selected based on clinical plausibility and pathophysiological relevance rather than univariable statistical significance, as univariable screening may inadvertently exclude important confounders. The area under the receiver operating characteristic curve (*AUC*) was calculated for PRD alone and for the multivariable model using 5-fold stratified cross-validation. Given the exploratory nature and limited statistical power (29 events, EPV=4.14, well below the recommended threshold of 10 [16]), these analyses should be considered hypothesis-generating.

Software: All analyses were performed using Python (version 3.13.12) with scikit-learn (version 1.6.0) for clustering, scipy (version 1.15.0) for statistical tests, statsmodels (version 0.14.0) for logistic regression, pandas (version 2.2.0) for data manipulation, and matplotlib (version 3.10.0) for visualization.

This study is reported in accordance with the STROBE cohort statement for observational studies.

The reporting of this study follows the STROBE guidelines for observational studies (Additional file 6: STROBE checklist).

## Results

### Study Population

A total of 108 patients were assessed for eligibility and 5 patients were excluded (Figure 1), in whom 4 lacked a valid PRD measurement and 1 had incomplete baseline variables. A valid PRD measurement was therefore available for 104 patients, of whom 103 with complete baseline data formed the final cohort for clustering analysis (mean age  $63.2 \pm 9.8$  years, 22.3% female, Table 1); the remaining 1 patient with PRD data but incomplete baseline variables was included only in the PRD correlation analyses ( $n=104$ ). The median PRD was 6.0 (IQR 3.0–10.0) bpm. Left atrial low voltage was present in 48/98 patients (49.0%) with available data. Among these 103 patients, 100 underwent ablation, 2 were lost to follow-up, and 98 had complete follow-up data for prognostic analysis (Figure 1).

### PRD Correlations with Clinical Variables

An independent two-sample *t*-test showed no significant difference in PRD between the two randomized measurement orders ( $P=0.44$ ), indicating that measurement order did not introduce

systematic bias. PRD showed significant correlations with multiple clinical indicators among the 104 patients with available PRD data (Figure 2, Additional file 1: Table S1). PRD was strongly correlated with heart rate ( $r=+0.570$ ,  $P<0.001$ ), 30-second heart rate ( $r=+0.524$ ,  $P<0.001$ ), neutrophil count ( $r=+0.445$ ,  $P<0.001$ ), WBC ( $r=+0.400$ ,  $P<0.001$ ), and LVEF ( $r=-0.382$ ,  $P<0.001$ ). PRD also correlated significantly with heart failure ( $r=+0.342$ ,  $P<0.001$ ), left ventricular end-systolic diameter ( $r=+0.312$ ,  $P=0.001$ ), and EHRA symptom score ( $r=+0.263$ ,  $P=0.007$ ).

#### Clustering Identified Five Distinct Phenotypes

K-means clustering ( $K=5$ , random\_state=42, silhouette score=0.196) identified five phenotypes, each representing a distinct clinical subgroup (Table 2, Figure 3):

Cluster 0 ( $n=16$ , 15.5%): "Markedly Elevated PRD with Preserved LVEF"

Markedly elevated PRD ( $17.2 \pm 8.5$  bpm), elevated heart rate ( $102.1 \pm 7.5$  bpm)

Preserved LVEF ( $62.3 \pm 3.8\%$ ), heart failure rate (37.5%)

NT-proBNP ( $561.7 \pm 264.1$  pg/ml)

Cluster 1 ( $n=45$ , 43.7%): "Young Low-PRD Phenotype"

Low PRD ( $5.6 \pm 3.8$  bpm), relatively young age ( $58.7 \pm 6.5$  years)

Normal LVEF ( $61.1 \pm 3.9\%$ ), low NT-proBNP ( $507.7 \pm 282.4$  pg/ml)

245 Largest cluster, representing relatively healthier AF patients

246

247 Cluster 2 ( $n=4$ , 3.9%): "Extreme Phenotype - Young with High PRD and Reduced LVEF"

248

249 Youngest age ( $43.2 \pm 9.1$  years), highest PRD ( $22.0 \pm 14.8$  bpm)

250 Lowest LVEF ( $38.9 \pm 8.1\%$ ), 100% heart failure

251 Highest inflammatory markers (WBC  $6.8 \pm 2.1 \times 10^9/L$ )

252 This extreme phenotype represents a noteworthy patient subgroup within our cohort: young AF

253 patients with severe ventricular dysfunction despite markedly elevated PRD

254

255 Cluster 3 ( $n=29$ , 28.2%): "Elderly Phenotype"

256

257 Oldest age ( $71.7 \pm 6.7$  years), low PRD ( $5.1 \pm 4.5$  bpm)

258 Elevated NT-proBNP ( $946.5 \pm 506.7$  pg/ml), highest stroke/TIA/thromboembolism rate (24.1%)

259 Representative of typical elderly AF patients

260

261 Cluster 4 ( $n=9$ , 8.7%): "High NT-proBNP Phenotype"

262

263 Highest NT-proBNP ( $2249 \pm 1103$  pg/ml), high heart rate and pulse rate

264 Moderate PRD ( $7.3 \pm 4.6$  bpm), 33.3% heart failure

265 Represents patients with elevated cardiac wall stress, as reflected by markedly elevated

266 NT-proBNP

The PCA visualization (Figure 3A) demonstrated that the five clusters occupied distinct regions in the principal component space, with C2 showing the greatest separation from other clusters. The radar chart (Figure 3B) further illustrated the unique feature profiles of each cluster, with C2 exhibiting the most deviant pattern across all dimensions.

#### Clustering Stability

To assess stability, we repeated clustering with 10 different random seeds. The mean Adjusted Rand Index (*ARI*) across 10 seeds was 0.724 (SD=0.12), and the mean reclassification rate was 11.4% (SD=5.4%), indicating moderate clustering reproducibility across random initialization seeds. The C2 core members ( $n=4$ ) remained together in 9/10 seeds; in seed 8, one edge patient (the one with LVEF=50%, the highest in C2) was reassigned to C0 while the other three stayed together. In 5/10 seeds (seeds 3–7), additional patients from neighboring clusters (C0, C1, or C3) were temporarily assigned to C2, but these were boundary cases reflecting centroid initialization sensitivity rather than instability of the C2 core phenotype (Additional file 2: Figure S1).

Sensitivity analyses: We conducted two sensitivity analyses by adding left atrial diameter (LAD) and the presence of left atrial low voltage separately to the 7-feature clustering model (8-feature models). Adding LAD ( $n=103$ ) yielded a lower silhouette score (0.159 vs 0.196) and Adjusted Rand Index (0.619 vs 0.724) than the 7-feature model, indicating reduced clustering stability. Adding left atrial low voltage ( $n=98$ ) yielded a marginally higher silhouette score (0.206 vs 0.196)

but a substantially lower Adjusted Rand Index (0.310 vs 0.724), indicating that this binary variable substantially reshuffled the cluster structure relative to the 7-feature model. In both augmentations, the C2 extreme phenotype (young age, high PRD, reduced LVEF) remained stable and intact. Based on these stability reductions, we retained the 7-feature model for parsimony and reproducibility.

Sensitivity analysis removing PRD: To evaluate whether PRD drives the cluster structure, we performed an additional sensitivity analysis by removing PRD from the feature set and repeating K-means clustering with the same parameters ( $k=5$ , seed=42, 6 features: HR, PR, LVEF, NT-proBNP, eGFR, age). The reclassification rate was 24.3% (25/103 patients), with an *ARI* of 0.608 between the full-model and PRD-excluded cluster assignments (Additional file 3: Figure S2). All 4 patients in the extreme C2 phenotype remained together in the same cluster even after removing PRD, suggesting that this subtype is stably defined by elevated PRD. It is further characterized by its constellation of associated features (young age, low LVEF, heart failure status).

#### Exploratory Prognostic Analysis

Among 98 patients with follow-up data, 29 (29.6%) experienced recurrence. In exploratory analyses, recurrence rates did not differ significantly across clusters ( $P=0.386$ ). The area under the ROC curve (*AUC*) for PRD alone predicting recurrence was 0.579. In a predefined multivariable model adjusting for clinically relevant variables (PRD, LAD, age, gender, LVEF, NT-proBNP, eGFR; 7 variables), the *AUC* was 0.534 (Additional file 4: Figure S3, Additional file 5: Table S2).

The minimal difference between the two models and the wide confidence intervals indicate that additional covariates conferred no statistically significant incremental prognostic information beyond PRD under cross-validation, consistent with the small sample and low events-per-variable (EPV=4.14). No variable achieved statistical significance in the multivariable model (PRD:  $OR=0.742$ , 95%  $CI$  0.438–1.257,  $P=0.268$ ). The primary value of this clustering approach lies in cross-sectional phenotyping rather than prognosis prediction (Additional file 5: Table S2).

## Discussion

### Main Findings

Key findings of the current study are: (1) PRD correlates significantly with heart rate, inflammatory markers, ventricular function, and heart failure history, supporting its clinical relevance as a bedside-available phenotypic marker; (2) unsupervised clustering identified five distinct AF phenotypes, including a previously unrecognized extreme subtype (C2: young age, high PRD, reduced LVEF, 100% heart failure); (3) PRD-derived clusters did not significantly associate with post-ablation recurrence in our exploratory analysis.

### Comparison with Prior Literature

Relatively few studies have investigated PRD in AF. The only prior clinical study of PRD in AF (Karadavut et al.) focused on exercise intolerance, not phenotyping [7]. Our study extends this

work by demonstrating PRD's utility for identifying distinct AF subgroups. More broadly, unsupervised phenotyping has already revealed clinically and prognostically distinct AF subgroups in large cohorts [17,18], establishing a strong rationale for incorporating previously overlooked bedside parameters such as PRD. Prior AF clustering studies have used electrical parameters (ECG markers) [4], imaging phenotypes [6], plasma biomarkers [5,13], and electronic health records [5], but none incorporated PRD. For example, Zhang et al. (2025) performed K-prototype clustering in the CABANA trial and identified phenotypes associated with different outcomes [9]; Soler-Espejo et al. (2026) identified 5 AF clusters in the MAFP-III cohort with markedly different cardiovascular mortality ( $aHR=6.82$  for the highest-risk cluster) [10]; and Gurnani et al. (2025) derived 5 AF phenotypes from AI-enhanced ECG [4] — yet none considered PRD. Similarly, Lin et al. (2026) developed an explainable ML model for stroke prediction in AF ( $AUC=0.91$ ) [19], and Bifulco et al. (2025) predicted post-ablation recurrence ( $AUC=0.80$ ) [20], highlighting the growing role of machine learning in AF risk stratification, though neither incorporated PRD. Our study demonstrates that adding PRD may reveal previously unrecognized subtypes (e.g., C2) that are invisible to existing phenotyping approaches.

#### PRD as a Phenotypic Marker

The primary value of this study lies in the phenotypic classification rather than prognosis prediction. The identification of distinct AF subtypes, particularly the extreme C2 phenotype, suggests that PRD captures clinically meaningful heterogeneity beyond traditional parameters.

While PRD alone may not predict recurrence, its integration into multidimensional phenotyping may inform future hypothesis-generating stratification approaches.

PRD's significant correlation with heart rate ( $r=+0.570$ ) is physiologically expected, as faster heart rates increase the proportion of beats with insufficient stroke volume to generate a peripheral pulse [7]. PRD also correlates with LVEF ( $r=-0.382$ ), heart failure ( $r=+0.342$ ), and inflammatory markers (WBC  $r=+0.400$ ). Because AF is characterized by intertwined electrical and structural remodeling [21] and by atrial cardiomyopathy [22], these correlations suggest that PRD may capture this multidimensional pathophysiology. This multidimensionality makes PRD a valuable addition to existing phenotypic markers.

PRD showed a weak but significant negative correlation with age ( $r=-0.194$ ,  $P=0.048$ ), indicating that younger AF patients tend to have higher pulse deficits. This finding is consistent with the C2 phenotype (youngest age, highest PRD). One plausible explanation is that age-related changes in atrioventricular (AV) node conduction modulate the ventricular response during AF. In younger patients, the AV node typically conducts a larger proportion of atrial impulses to the ventricles, resulting in faster and more irregular ventricular rates, which in turn increase the proportion of beats with insufficient stroke volume to generate a palpable pulse. Conversely, age-related AV node conduction slowing in older patients limits ventricular rate, leading to fewer pulse deficits and a lower PRD. This age-PRD relationship further underscores the heterogeneity of AF across age groups and suggests that the clinical significance of PRD may differ between young and elderly patients.

PRD may complement existing biomarker-based assessment in AF. Our correlation analysis revealed that PRD correlated with EHRA symptom score ( $r=+0.263$ ,  $P=0.007$ ) but showed no significant correlation with NT-proBNP ( $r=-0.062$ ,  $P=0.536$ ). This divergence suggests that PRD and NT-proBNP may reflect different dimensions of hemodynamic burden in AF, with PRD more closely associated with patient-reported symptoms. In clinical practice, relying solely on NT-proBNP levels may not fully capture the symptomatic burden in some patients; PRD, as a rapid bedside measurement, could serve as a complementary tool to identify patients who warrant closer evaluation—particularly those with significant symptoms but unremarkable biomarker profiles. This is especially relevant for the C0 phenotype, where patients exhibited markedly elevated PRD despite preserved LVEF, suggesting that PRD may identify hemodynamic impairment even when conventional echocardiographic parameters remain within normal ranges.

#### The Distinct C2 Phenotype

The most distinctive finding is Cluster 2 ( $n=4$ ): remarkably young (43.2 years), highest PRD (22.0 bpm), lowest LVEF (38.9%), and 100% heart failure. This phenotype may reflect a severe form of hemodynamic impairment related to irregular RR intervals [22,23]; however, the cross-sectional design precludes determining whether irregular RR intervals lead to reduced LVEF or whether pre-existing cardiomyopathy rendered these patients more susceptible to a pronounced pulse deficit. The coexistence of high PRD and reduced LVEF suggests that severe pulse deficit may be associated with ventricular dysfunction, independent of rate alone.

399

400 Several mechanisms may explain this phenotype: (1) loss of atrial kick: in AF, loss of atrial  
401 contraction reduces cardiac output by 15–30%, and the resulting hemodynamic perturbation  
402 engages the cardiac baroreflex within a closed-loop system, wherein baroreflex suppression  
403 further compromises ventricular function in already compromised ventricles [23]; (2) RR interval  
404 irregularity: AF-induced RR chaos further impairs ventricular filling and increases myocardial  
405 oxygen demand, as reviewed in [21] and supported by prognostic data in HF+AF cohorts [24]; (3)  
406 neurohormonal activation: chronic AF activates the renin-angiotensin system, promoting  
407 ventricular remodeling, as reviewed in [22].

408

409 The identification of the C2 phenotype — young AF patients with high PRD and reduced LVEF  
410 — raises the hypothesis that early rhythm control might attenuate further ventricular deterioration  
411 in this subgroup; this remains to be tested prospectively, as our cross-sectional design cannot  
412 establish therapeutic benefit. C2's core stability (4/4 members intact in 9/10 seeds) suggests it is  
413 not a clustering artifact but a reproducible phenotype, although boundary patients from  
414 neighboring clusters were occasionally absorbed into C2 under different initialization seeds. The  
415 persistence of the C2 phenotype after PRD removal from the clustering features (all 4 patients  
416 remained together) suggests that it is not merely a univariate artifact driven solely by PRD  
417 elevation, but rather a stable, multidimensional patient subgroup anchored by the concurrent  
418 presence of young age, severely reduced LVEF, and heart failure — with PRD potentially  
419 reflecting this underlying hemodynamic burden. Although C2 had the most severe hemodynamic  
420 phenotype, NT-proBNP was not the highest, suggesting that PRD may reflect acute hemodynamic

compromise beyond biomarker elevation.

Notably, although clusters C2 and C4 had nearly identical heart rates (103.75 vs 104.11 bpm), their PRD differed markedly (22.00 vs 7.33 bpm), reflecting a fundamental difference in the efficiency of cardiac output. This contrast suggests that C2 may represent a tachycardia-mediated cardiomyopathy-like phenotype [22], in which rapid ventricular rates generate a high proportion of hemodynamically ineffective beats, whereas C4 reflects preserved stroke volume despite a rapid rate. Given the cross-sectional design, the reversibility of the reduced LVEF in C2 after rate control could not be assessed and warrants prospective validation.

#### Exploratory Prognostic Findings

Applying unsupervised clustering to derive AF phenotypes and subsequently testing their association with clinical outcomes is an established analytical strategy [17,18]. In line with this approach, we conducted an exploratory analysis of whether the PRD-derived phenotypes were associated with post-ablation recurrence. We did not observe a significant association ( $P=0.386$ ). This null finding should not be over-interpreted: the phenotypes were derived to capture baseline clinical heterogeneity rather than to define a recurrence substrate, and prior unsupervised AF phenotyping studies have been primarily linked to hard cardiovascular endpoints rather than post-ablation recurrence [17,18]. The modest sample size ( $n=103$ ) and limited number of recurrence events ( $n=29$ ) further reduced statistical power to detect modest effect sizes. Larger studies are needed to determine whether these phenotypes carry prognostic implications.

443

444 Several additional factors may contribute to the null prognostic finding: (1) treatment effect  
445 confounding, as high-PRD patients (especially C2) likely received more aggressive management;  
446 (2) the multifactorial nature of AF recurrence, which depends on atrial fibrosis burden [6], left  
447 atrial size, and operator experience beyond baseline PRD [25]; (3) the universal 3-month  
448 antiarrhythmic drug use post-ablation, which may have suppressed recurrence and biased results  
449 toward the null; and (4) the relatively short 6-month follow-up, which may be insufficient to  
450 capture late recurrences; and (5) the primary aim was phenotypic discovery and exploration of  
451 clinical significance rather than recurrence prediction, such that a null association does not detract  
452 from the clinical value of these subtypes.

453

454

455

456

457

458 Future Directions

459 The recent PFA-SHAM and ADVENT-LTO trials have established the efficacy and durability of  
460 pulsed field ablation for AF [26,27], supporting future validation of PRD-guided early intervention  
461 in high-PRD phenotypes. Continuous PRD assessment could be enabled by wearable ECG-PPG  
462 devices—extending Zuo et al.’s demonstration of PPG-derived AF-burden accuracy ( $r=0.999$ ) [28]  
463 and Ding et al.’s cross-modal PPG approach [29]—to longitudinally track hemodynamic burden,  
464 although technical validation is required.

## Limitations

This study has several limitations. First, the single-center design and modest sample size limit generalizability and statistical power. With only 29 recurrence events and 7 predictor variables, the study was substantially underpowered (EPV=4.14, well below the recommended threshold of 10 [16]), and the prognostic analyses were therefore exploratory and hypothesis-generating. This limited power may explain why traditional predictors (LAD, NT-proBNP, LVEF) were also not significantly associated with recurrence in our cohort, contrary to prior literature [6,20]. Emerging evidence suggests that AF-related cerebral perfusion impairment [30] and symptom burden networks [31] may represent alternative prognostic dimensions not captured by recurrence alone. Second, the silhouette score of 0.196 at  $K=5$  was relatively low, indicating considerable inter-cluster overlap, which is consistent with a continuous biological continuum of AF phenotypes rather than discrete subtypes. The extremely unequal cluster sizes, particularly the small C2 subgroup ( $n=4$ ), further warrant cautious interpretation of subgroup-specific findings; however, the stability of C2's core members across 9/10 random seeds mitigates concerns about a random artifact. Third, the cross-sectional baseline design precludes causal inference; the observed associations between PRD, inflammatory markers, and ventricular function represent correlation rather than causation. In particular, systematic post-ablation LVEF reassessment was not available in this cohort, so the reversibility of the reduced LVEF in C2 and the causal direction of this association cannot be determined. Fourth, PRD was acquired by a single investigator using manual auscultation and palpation. Although measurement order was randomized and showed

no systematic bias ( $P=0.44$ ), intra- and inter-rater reproducibility were not formally assessed. No standardized protocol for PRD measurement currently exists, which may limit cross-center generalizability. Fifth, LAD and left atrial low voltage were deliberately excluded from clustering (see Results for sensitivity analyses), and the universal 3-month post-ablation amiodarone use may have suppressed recurrence, biasing prognostic analyses toward the null. Finally, the 6-month follow-up may be insufficient to capture late recurrences.

## Conclusions

PRD is a bedside-available phenotypic marker that captures clinically meaningful phenotypic heterogeneity, identifying distinct AF subgroups including a reproducible extreme subtype (C2) characterized by young age, high PRD, reduced LVEF, and 100% heart failure. In exploratory analyses, PRD-derived clusters were not significantly associated with post-ablation recurrence. Future studies should validate these phenotypes in larger cohorts and explore whether PRD-guided therapy improves outcomes.

## List of abbreviations

AF: atrial fibrillation; AI: artificial intelligence; ANOVA: analysis of variance; ARI: adjusted Rand index; AUC: area under the curve; AV: atrioventricular; BMI: body mass index; CI: confidence interval; ECG: electrocardiogram; eGFR: estimated glomerular filtration rate; EHRA: European Heart Rhythm Association; EPV: events per variable; HF: heart failure; IQR: interquartile range; LAD: left atrial diameter; LVA: low-voltage area; LVDs: left ventricular end-systolic diameter; LVEDD: left ventricular end-diastolic diameter; LVEF: left ventricular ejection fraction; ML:

509 machine learning; NT-proBNP: N-terminal pro-B-type natriuretic peptide; OR: odds ratio; PCA:  
510 principal component analysis; PFA: pulsed field ablation; PPG: photoplethysmography; PRD: pulse  
511 rate deficit; RAAS: renin-angiotensin-aldosterone system; ROC: receiver operating characteristic;  
512 RR: R-R interval; SD: standard deviation; STROBE: Strengthening the Reporting of Observational  
513 Studies in Epidemiology; TIA: transient ischemic attack; WBC: white blood cell

514

#### 515 **Declarations**

516

#### 517 **Ethics approval and consent to participate**

518 The study protocol was approved by the Institutional Review Board of the First Affiliated Hospital  
519 of Nanjing Medical University. The study was conducted in accordance with the Declaration of  
520 Helsinki. Written informed consent was obtained from all participants prior to enrollment.

521

#### 522 **Consent for publication**

523 Not applicable.

524

#### 525 **Availability of data and materials**

526 The datasets used and analysed during the current study are not publicly available because they  
527 contain clinical information that could compromise the privacy of the participants and the  
528 confidentiality conditions of the ethics approval. De-identified data are available from the  
529 corresponding author on reasonable request, subject to approval by the Institutional Review  
530 Board of the First Affiliated Hospital of Nanjing Medical University.

531

532 **Competing interests**

533 The authors declare that they have no competing interests.

534

535 **Funding**

536 This research received no external funding.

537

538 **Authors' contributions**

539 MC, XX and HL conceived and designed the study. KL performed the PRD measurements (heart  
540 rate and pulse rate). XX collected the clinical data. XX performed the statistical analysis. XX  
541 drafted the manuscript. CC contributed to the literature search and assisted in revising the  
542 manuscript. HL and MC critically revised the manuscript for important intellectual content. All  
543 authors read and approved the final manuscript.

544

545

546 **References**

- 547 [1] Chugh SS, Havmoeller R, Narayanan K, Singh D, Rienstra M, Benjamin EJ, Gillum RF, Kim YH,  
548 McAnulty JH, Zheng ZJ, Forouzanfar MH, Naghavi M, Mensah GA, Ezzati M, Murray CJ. Worldwide  
549 epidemiology of atrial fibrillation: a Global Burden of Disease 2010 Study. Circulation.  
550 2014;129(8):837-847. doi: 10.1161/CIRCULATIONAHA.113.005119.
- 551 [2] Van Gelder IC, Rienstra M, Bunting KV, et al. 2024 ESC Guidelines for the management of  
552 atrial fibrillation developed in collaboration with the European Association for Cardio-Thoracic

553 Surgery (EACTS). *Eur Heart J*. 2024;45(36):3314-3414. doi: 10.1093/eurheartj/ehae176. Erratum  
554 in: *Eur Heart J*. 2025;46(41):4349. doi: 10.1093/eurheartj/ehaf306. PMID: 39210723.

555 [3] Tzeis S, Gerstenfeld EP, Kalman J, et al. 2024 European Heart Rhythm Association/Heart  
556 Rhythm Society/Asia Pacific Heart Rhythm Society/Latin American Heart Rhythm Society expert  
557 consensus statement on catheter and surgical ablation of atrial fibrillation. *Europace*.  
558 2024;26(4):euae043. doi: 10.1093/europace/euae043.

559 [4] Gurnani M, Patlatzoglou K, Barker J, et al. Deriving novel atrial fibrillation phenotypes using  
560 a tree-based artificial intelligence-enhanced electrocardiography approach. *NPJ Digit Med*.  
561 2025;8(1):779. doi: 10.1038/s41746-025-02159-z.

562 [5] Zhang Z, Yoshimura H, Acosta-Mena D, et al. Clustering of electronic health records in atrial  
563 fibrillation patients and impact on prognosis and patient trajectories: a UK linked-dataset study.  
564 *Eur Heart J Digit Health*. 2025;6(4):797-810. doi: 10.1093/ehjdh/ztaf032.

565 [6] Marrouche NF, Wilber D, Hindricks G, et al. Association of atrial tissue fibrosis identified by  
566 delayed enhancement MRI and atrial fibrillation catheter ablation: the DECAAF study. *JAMA*.  
567 2014;311(5):498-506. doi: 10.1001/jama.2014.3.

568 [7] Karadavut S, Altintop I. Pulse deficit in atrial fibrillation – a different perspective on rhythm or  
569 rate control strategy. *Kardiol Pol*. 2021;79(11):1231-1238. doi: 10.33963/KP.a2021.0107.

570 [8] Wijesurendra RS, Casadei B. Atrial fibrillation: effects beyond the atrium? *Cardiovasc Res*.  
571 2015;105(3):238-247. doi: 10.1093/cvr/cvv001.

572 [9] Zhang J, Wang L, Li Y, et al. Atrial fibrillation phenotypes identified through cluster analysis  
573 in the CABANA study. *Int J Cardiol*. 2025;438:133606. doi: 10.1016/j.ijcard.2025.133606.

574 [10] Soler-Espejo E, Ramos-Bratos MP, Chen Y, et al. Phenotype-guided risk stratification in

575 anticoagulated atrial fibrillation patients: cluster analysis of the Murcia AF Project III (MAFP-III).  
576 Heart Rhythm. 2026;23(3):563-574. doi: 10.1016/j.hrthm.2025.09.003.

577 [11] Yang B, Jiang C, Lin Y, et al. STABLE-SR (electrophysiological substrate ablation in the left  
578 atrium during sinus rhythm) for the treatment of nonparoxysmal atrial fibrillation: a prospective,  
579 multicentre randomized clinical trial. Circ Arrhythm Electrophysiol. 2017;10(11):e005405. doi:  
580 10.1161/CIRCEP.117.005405.

581 [12] Yang G, Zheng L, Jiang C, et al. Circumferential pulmonary vein isolation plus  
582 low-voltage-area modification in persistent atrial fibrillation: the STABLE-SR-II trial. JACC Clin  
583 Electrophysiol. 2022;8(7):882-891. doi: 10.1016/j.jacep.2022.03.012.

584 [13] Bresso E, Lacomblez C, Duarte K, et al. Unsupervised machine learning for cardiovascular  
585 disease: a framework for future studies. Eur J Heart Fail. 2025;27(12):3366-3377. doi:  
586 10.1002/ejhf.70076.

587 [14] Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster  
588 analysis. J Comput Appl Math. 1987;20:53-65. doi: 10.1016/0377-0427(87)90125-7.

589 [15] Hubert L, Arabie P. Comparing partitions. J Classif. 1985;2(1):193-218. doi:  
590 10.1007/BF01908075.

591 [16] Peduzzi P, Concato J, Kemper E, et al. A simulation study of the number of events per  
592 variable in logistic regression analysis. J Clin Epidemiol. 1996;49(12):1373-1379. doi:  
593 10.1016/s0895-4356(96)00236-3.

594 [17] Inohara T, Shrader P, Pieper K, et al. Association of atrial fibrillation clinical phenotypes with  
595 treatment patterns and outcomes: a multicenter registry study. JAMA Cardiol. 2018;3(1):54-63.  
596 doi: 10.1001/jamacardio.2017.4665.

597 [18] Vitolo M, Proietti M, Shantsila A, et al. Clinical phenotype classification of atrial fibrillation  
 598 patients using cluster analysis and associations with trial-adjudicated outcomes. *Biomedicines*.  
 599 2021;9(7):843. doi: 10.3390/biomedicines9070843.

600 [19] Lin JC, Chang CM, Pan HY, et al. Interpretable machine learning models for stroke risk  
 601 prediction in patients with newly diagnosed atrial fibrillation. *NPJ Digit Med*. 2026;9(1):289. doi:  
 602 10.1038/s41746-026-02470-3.

603 [20] Bifulco SF, Magoon MJ, Chahine Y, et al. Predicting arrhythmia recurrence post-ablation in  
 604 atrial fibrillation using explainable machine learning. *Commun Med (Lond)*. 2025;5(1):421. doi:  
 605 10.1038/s43856-025-01058-4.

606 [21] Boehmer AA, Ninni S, Heijman J, et al. The clinical pathophysiology of atrial fibrillation:  
 607 outstanding questions from bedside to bench and back. *Physiol Rev*. 2026;106(3):1055-1121. doi:  
 608 10.1152/physrev.00020.2025.

609 [22] Fabritz L, Hatem S, Schnabel R, et al. Arrhythmia-induced cardiomyopathy: focus on atrial  
 610 fibrillation. *Nat Rev Cardiol*. 2026;23(1):1-19. doi: 10.1038/s41569-025-01195-2.

611 [23] Adeodu O, Gee M, Mahmoudi B, et al. Short-term hemodynamic effects of atrial fibrillation  
 612 in a closed-loop human cardiac-baroreflex system. *PLoS One*. 2025;20(10):e0334086. doi:  
 613 10.1371/journal.pone.0334086.

614 [24] Wang C, Chen H, Ye J, et al. Heart-rate-derived variability measures and in-hospital mortality  
 615 in critically ill patients with heart failure and atrial fibrillation: insights from the MIMIC-IV  
 616 database. *BMC Cardiovasc Disord*. 2026;26(1):487. doi: 10.1186/s12872-026-05888-5.

617 [25] Anselmino M, Battaglia A, Matta M, et al. Machine learning prediction of outcome following  
 618 pulsed-field atrial fibrillation ablation. *Europace*. 2026;28(5):euag053. doi:

10.1093/europace/euag053.

[26] Osmancik P, Neuzil P, Hozmanova J, et al. Pulsed field ablation versus sham to treat atrial fibrillation: the PFA-SHAM randomized clinical trial. *Circulation*. 2026;153(25):1984-1998. doi: 10.1161/CIRCULATIONAHA.126.079484.

[27] Reddy VY, Gerstenfeld EP, Mountantonakis SE, et al. Pulsed field ablation versus conventional thermal ablation for paroxysmal atrial fibrillation: 4-year outcomes in the ADVENT-LTO study. *Nat Med*. 2026;32(4):1444-1453. doi: 10.1038/s41591-026-04246-4.

[28] Zuo S, Zhou L, Feng H, et al. Validation of smartwatches integrated with photoplethysmography for continuous evaluation of atrial fibrillation burden. *JACC Clin Electrophysiol*. 2026;12(6):589-598. doi: 10.1016/j.jacep.2026.03.036.

[29] Ding Z, Hu Y, Li Z, et al. AI modeling photoplethysmography to electrocardiography useful for predicting cardiovascular disease. *NPJ Digit Med*. 2025;9(1):61. doi: 10.1038/s41746-025-02228-3.

[30] Borch P, Aagaard-Nilsen S, Stokke MK, et al. Effects of atrial fibrillation on cerebral perfusion and cognitive function: a systematic review. *Int J Cardiol Heart Vasc*. 2026;63:101891. doi: 10.1016/j.ijcha.2026.101891.

[31] Sun D, Yang X, Li H, et al. Core and bridging symptoms in patients with atrial fibrillation: a network analysis. *Front Cardiovasc Med*. 2025;12:1617872. doi: 10.3389/fcvm.2025.1617872.

## Tables

Table 1. Baseline characteristics of 103 patients undergoing clustering analysis

| Variable                                               | Total (N=103)       |
|--------------------------------------------------------|---------------------|
| <b>Demographics</b>                                    |                     |
| Age (years)                                            | 63.2 ± 9.8          |
| Female gender                                          | 23 (22.3%)          |
| BMI (kg/m <sup>2</sup> )                               | 25.7 ± 3.6          |
| <b>AF characteristics</b>                              |                     |
| AF duration (months)                                   | 18.0 (3.2–57.0)     |
| EHRA score                                             | 2.0 ± 1.3           |
| <b>PRD parameters</b>                                  |                     |
| PRD (bpm)                                              | 6.0 (3.0–10.0)      |
| Heart rate (bpm)                                       | 85.4 ± 14.1         |
| Pulse rate (bpm)                                       | 79.5 ± 12.5         |
| <b>Echocardiography</b>                                |                     |
| LVEF (%)                                               | 59.9 ± 6.4          |
| LVEDD (mm)                                             | 48.4 ± 3.8          |
| LVDs (mm)                                              | 33.0 ± 4.2          |
| LAD (mm)                                               | 44.3 ± 3.5          |
| Left atrial low voltage, n/N (%)                       | 48/98 (49.0%)       |
| <b>Laboratory</b>                                      |                     |
| NT-proBNP (pg/ml)                                      | 594.6 (375.2–993.8) |
| eGFR (ml/min/1.73m <sup>2</sup> )                      | 85.7 ± 16.5         |
| WBC (×10 <sup>9</sup> /L)                              | 5.8 ± 1.5           |
| Neutrophils (×10 <sup>9</sup> /L)                      | 3.3 ± 1.1           |
| Hemoglobin (g/L)                                       | 145.3 ± 16.5        |
| Albumin (g/L)                                          | 39.7 ± 3.0          |
| Indirect bilirubin (μmol/L)                            | 9.8 ± 4.4           |
| <b>Comorbidities</b>                                   |                     |
| Hypertension                                           | 59 (57.3%)          |
| Diabetes mellitus                                      | 14 (13.6%)          |
| Heart failure                                          | 32 (31.1%)          |
| Prior stroke/TIA/thromboembolism                       | 15 (14.6%)          |
| Coronary artery disease or peripheral vascular disease | 19 (18.4%)          |

Data are mean ± SD, median (IQR), or n (%).

\*Left atrial low voltage data were available for 98 patients; the denominator reflects available data.

647

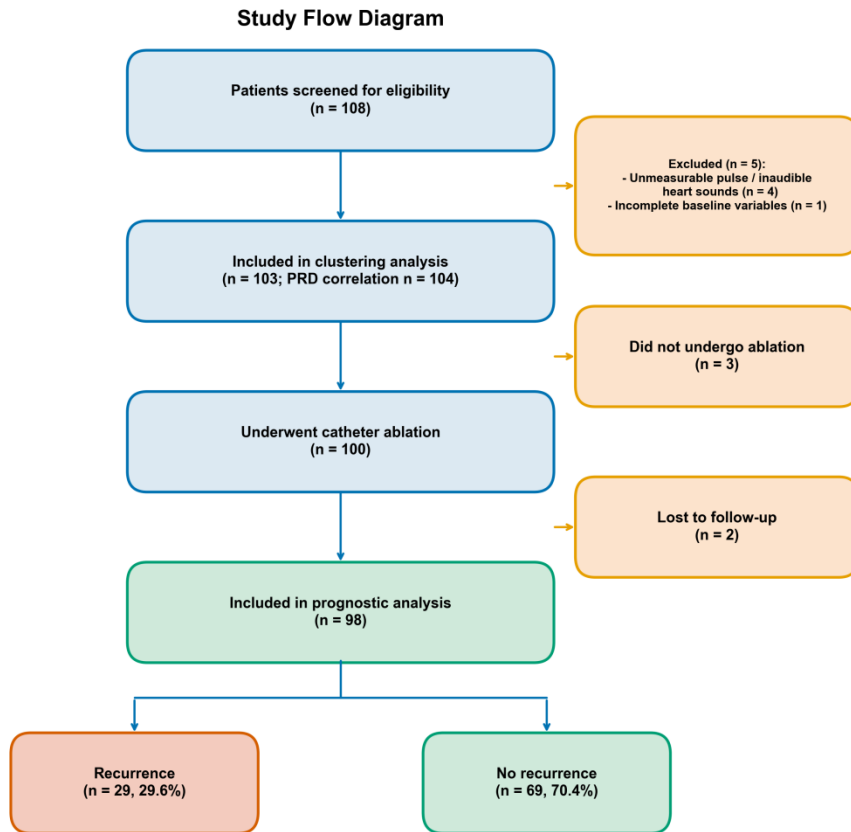
648 Table 2. Clinical characteristics of 5 clusters identified by K-means clustering

649

| Variable                                     | C0 (n=16)   |   | C1 (n=45)   |   | C2 (n=4)    |   | C3 (n=29)   |   | C4 (n=9)    |   | P value |
|----------------------------------------------|-------------|---|-------------|---|-------------|---|-------------|---|-------------|---|---------|
| Demographics                                 |             |   |             |   |             |   |             |   |             |   |         |
| Age (years)                                  | 61.88       | ± | 58.69       | ± | 43.25       | ± | 71.66       | ± | 70.11       | ± | <0.001  |
|                                              | 8.34        |   | 6.51        |   | 8.26        |   | 6.73        |   | 6.21        |   |         |
| Female gender, n (%)                         | 4(25%)      |   | 9(20%)      |   | 0(0%)       |   | 6(21%)      |   | 4(44%)      |   | 0.414   |
| BMI (kg/m²)                                  | 26.62       | ± | 26.17       | ± | 26.77       | ± | 24.72       | ± | 24.66       | ± | 0.271   |
|                                              | 3.28        |   | 3.06        |   | 2.17        |   | 4.43        |   | 3.40        |   |         |
| AF characteristics                           |             |   |             |   |             |   |             |   |             |   |         |
| AF duration (months)                         | 46.66       | ± | 40.88       | ± | 71.50       | ± | 26.49       | ± | 33.39       | ± | 0.525   |
|                                              | 53.58       |   | 62.21       |   | 88.23       |   | 41.87       |   | 51.83       |   |         |
| EHRA score                                   | 2.00 ± 1.46 |   | 1.73 ± 1.18 |   | 3.75 ± 0.50 |   | 2.00 ± 1.28 |   | 2.44 ± 1.33 |   | 0.033   |
| PRD parameters                               |             |   |             |   |             |   |             |   |             |   |         |
| PRD (bpm)                                    | 17.19       | ± | 5.60 ± 3.75 |   | 22.00       | ± | 5.07 ± 4.54 |   | 7.33 ± 4.58 |   | <0.001  |
|                                              | 8.54        |   |             |   | 14.76       |   |             |   |             |   |         |
| Heart rate (bpm)                             | 102.06      | ± | 81.31       | ± | 103.75      | ± | 74.21       | ± | 104.11      | ± | <0.001  |
|                                              | 7.50        |   | 9.05        |   | 4.79        |   | 7.80        |   | 8.77        |   |         |
| Pulse rate (bpm)                             | 87.38       | ± | 76.51       | ± | 81.75       | ± | 72.72       | ± | 101.22      | ± | <0.001  |
|                                              | 10.98       |   | 8.43        |   | 13.07       |   | 8.23        |   | 13.61       |   |         |
| Echocardiography                             |             |   |             |   |             |   |             |   |             |   |         |
| LAD (mm)                                     | 44.38       | ± | 43.64       | ± | 46.00       | ± | 44.66       | ± | 45.22       | ± | 0.480   |
|                                              | 3.34        |   | 3.16        |   | 2.45        |   | 3.64        |   | 4.79        |   |         |
| LVEDD (mm)                                   | 48.06       | ± | 48.60       | ± | 56.25       | ± | 47.86       | ± | 46.78       | ± | <0.001  |
|                                              | 2.84        |   | 3.11        |   | 7.09        |   | 3.81        |   | 3.23        |   |         |
| LVDs (mm)                                    | 31.88       | ± | 32.69       | ± | 45.60       | ± | 32.66       | ± | 31.89       | ± | <0.001  |
|                                              | 2.28        |   | 2.91        |   | 7.76        |   | 4.02        |   | 2.98        |   |         |
| LVEF (%)                                     | 62.30       | ± | 61.05       | ± | 38.92       | ± | 59.70       | ± | 60.00       | ± | <0.001  |
|                                              | 3.80        |   | 3.93        |   | 8.08        |   | 6.08        |   | 4.21        |   |         |
| Clinical scores                              |             |   |             |   |             |   |             |   |             |   |         |
| CHA <sub>2</sub> DS <sub>2</sub> -VASc score | 2.31 ± 1.89 |   | 1.51 ± 1.34 |   | 1.50 ± 0.58 |   | 3.24 ± 1.70 |   | 3.44 ± 1.74 |   | <0.001  |
| APPLE score                                  | 2.19 ± 0.91 |   | 1.86 ± 0.73 |   | 2.25 ± 0.50 |   | 3.14 ± 0.92 |   | 3.67 ± 0.71 |   | <0.001  |
| Laboratory                                   |             |   |             |   |             |   |             |   |             |   |         |
| NT-proBNP (pg/ml)                            | 561.73      | ± | 507.68      | ± | 531.77      | ± | 946.51      | ± | 2249.22     | ± | <0.001  |
|                                              | 264.12      |   | 282.40      |   | 341.56      |   | 506.72      |   | 1102.5      |   |         |
| eGFR (ml/min/1.73m²)                         | 85.25       | ± | 94.76       | ± | 106.25      | ± | 75.48       | ± | 65.00       | ± | <0.001  |
|                                              | 14.25       |   | 10.38       |   | 10.40       |   | 11.39       |   | 21.82       |   |         |
| WBC (×10 <sup>9</sup> /L)                    | 6.59 ± 1.38 |   | 5.64 ± 1.52 |   | 6.82 ± 2.07 |   | 5.54 ± 1.33 |   | 5.60 ± 0.90 |   | 0.077   |
| Neutrophils (×10 <sup>9</sup> /L)            | 3.93 ± 1.06 |   | 3.14 ± 0.99 |   | 4.18 ± 2.18 |   | 3.26 ± 0.95 |   | 3.27 ± 0.64 |   | 0.048   |
| Hemoglobin (g/L)                             | 152.38      | ± | 147.51      | ± | 146.00      | ± | 145.07      | ± | 122.56      | ± | <0.001  |

|                                                               |                    |                    |                    |                    |                    |       |
|---------------------------------------------------------------|--------------------|--------------------|--------------------|--------------------|--------------------|-------|
|                                                               | 14.38              | 13.32              | 14.81              | 14.81              | 23.59              |       |
| Platelet count ( $\times 10^9/L$ )                            | 195.69 $\pm$ 48.25 | 180.04 $\pm$ 57.30 | 168.00 $\pm$ 19.51 | 154.90 $\pm$ 41.36 | 190.89 $\pm$ 70.80 | 0.098 |
| <b>Comorbidities, n (%)</b>                                   |                    |                    |                    |                    |                    |       |
| Heart failure, n (%)                                          | 6(37.5%)           | 9(20%)             | 4(100%)            | 10(34%)            | 3(33%)             | 0.018 |
| Hypertension, n (%)                                           | 12(75%)            | 24(53%)            | 1(25%)             | 16(55%)            | 6(67%)             | 0.352 |
| Diabetes mellitus, n (%)                                      | 2(12%)             | 5(11%)             | 0(0%)              | 5(17%)             | 2(22%)             | 0.776 |
| <b>Prior stroke/TIA/thromboembolism, n (%)</b>                |                    |                    |                    |                    |                    |       |
| Coronary artery disease or peripheral vascular disease, n (%) | 1(6%)              | 5(11%)             | 0(0%)              | 7(24%)             | 2(22%)             | 0.335 |
|                                                               | 4(25%)             | 4(9%)              | 1(25%)             | 7(24%)             | 3(33%)             | 0.262 |

650 Data are mean $\pm$ SD or n (%). P values from one-way ANOVA (continuous variables) or chi-square/Fisher's exact test (categorical  
651 variables). For NT-proBNP and APPLE in C1, data were available for 44 patients (one missing value).  
652  
653 Figure Legends  
654  
655 Figure 1. Study flow diagram. Flow of patients from screening (n=108) to clustering analysis  
656 (n=103) to prognostic analysis (n=98). A total of 104 patients had valid PRD data and were  
657 included in the PRD correlation analyses; one patient with valid PRD data had incomplete  
658 baseline variables and was included in the correlation analyses but excluded from the clustering  
659 analysis.



PRD = pulse rate deficit = [auscultated heart rate – palpable pulse rate]  
 PRD correlation analyses included 104 patients. One patient with valid PRD data had incomplete baseline variables and was included in the correlation analyses but excluded from clustering.

660

661

662

663 Figure 2. Correlation between PRD and key clinical variables. Bar plot showing Pearson

664 correlation coefficients (r) between PRD and clinical variables, sorted by |r|. Red bars indicate

665 positive correlation, blue bars indicate negative correlation. \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .

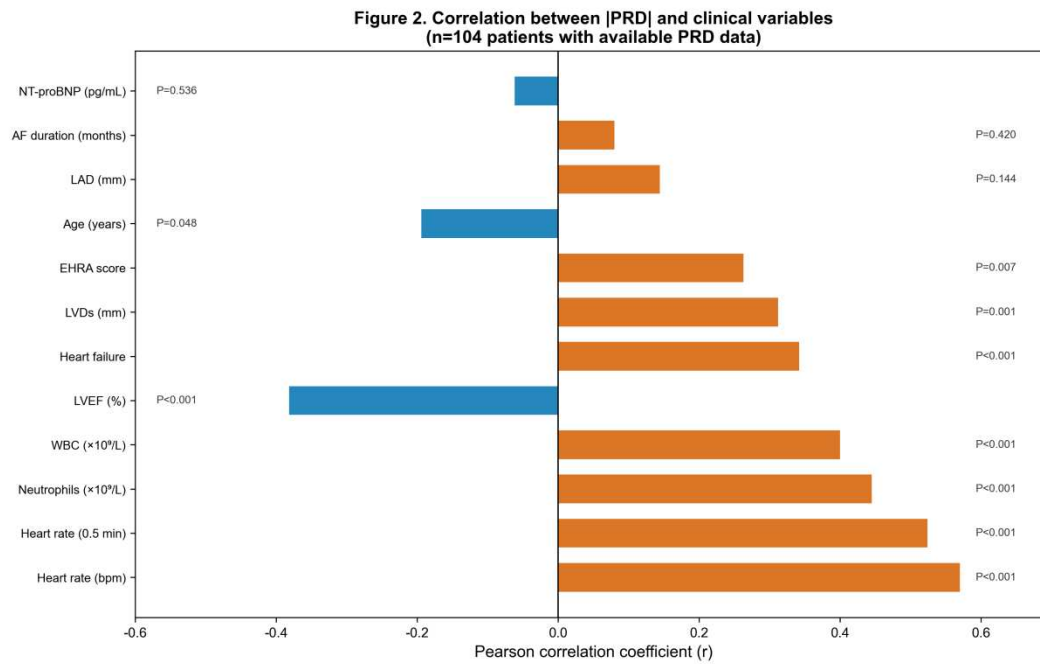
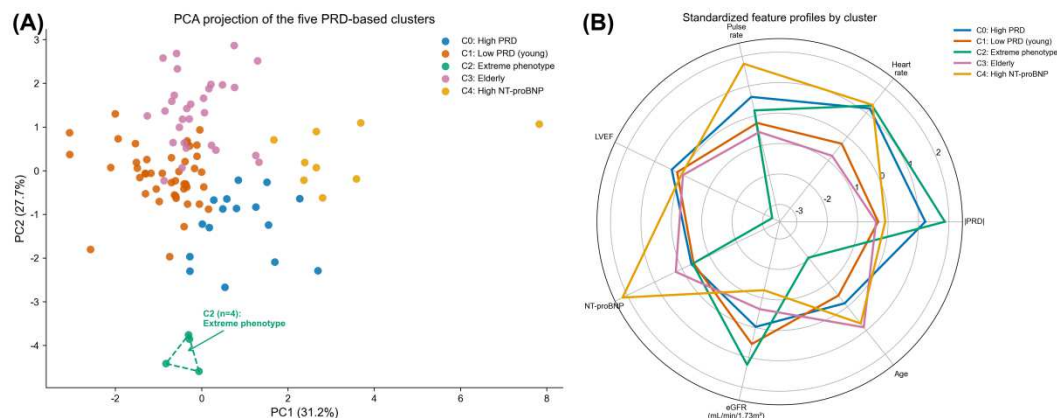


Figure 3. Cluster visualization. (A) Principal component analysis (PCA) scatter plot showing the distribution of 103 patients across 5 clusters in the first two principal components. Each color represents a different cluster. C2 (extreme phenotype) shows the greatest separation from other clusters. PC1 explained 31.2% and PC2 explained 27.7% of the variance; cumulative 58.9%. (B) Radar chart showing standardized feature profiles (z-scores) for each cluster across the 7 clustering variables. C2 exhibits the most deviant pattern, characterized by markedly elevated PRD, lowest LVEF, and elevated heart rate.



## **Additional files**

Additional file 1 — File name: Additional\_file\_1\_Table\_S1\_PRD\_correlation.xlsx. File format: .xlsx.

Title of data: Complete PRD correlation matrix (Table S1). Description of data: Pearson correlation coefficients ( $r$ ) and P values between PRD and all clinical variables tested; left atrial low voltage data were available for 98 of 103 patients.

Additional file 2 — File name: Additional\_file\_2\_Figure\_S1\_clustering\_stability.png. File format: .png. Title of data: Clustering stability analysis (Figure S1). Description of data: Heatmap showing patient assignments across 10 different random seeds (seed 1–10) compared to the reference clustering (seed 42). Patients are sorted by their seed-42 cluster assignment. C2 core members ( $n=4$ ) remained together in 9/10 seeds; in seed 8, one edge patient (LVEF=50%, the highest in C2) was reassigned to C0 while the other three C2 patients stayed together. In several seeds (3–7), additional patients from C0, C1, or C3 were temporarily assigned to C2, reflecting K-means sensitivity to centroid initialization at cluster boundaries rather than instability of the C2 core phenotype.

Additional file 3 — File name: Additional\_file\_3\_Figure\_S2\_PRD\_reclassification.png. File format: .png. Title of data: PRD reclassification sensitivity analysis (Figure S2). Description of data: Confusion matrix comparing cluster assignments with PRD (7 features) versus without PRD (6 features) in K-means clustering ( $k=5$ , seed=42). Diagonal elements represent patients who remained in the same cluster; off-diagonal elements represent reclassified patients.

699   Reclassification rate=24.3%, ARI=0.608. All 4 C2 patients remained together in the same cluster  
700   after PRD removal.

701

702   Additional file 4 — File name: Additional\_file\_4\_Figure\_S3\_ROC\_curve.png. File format: .png.

703   Title of data: ROC curve for PRD predicting post-ablation recurrence (Figure S3). Description of

704   data: Receiver operating characteristic curves; both models used 5-fold stratified cross-validation.

705   Blue line: PRD alone (AUC=0.579, 95% CI 0.456–0.700). Red line: multivariable model (AUC=0.534,

706   95% CI 0.374–0.625). Dashed line: reference. DeLong test P=0.546 (not statistically significant).

707   X-axis: 1 – Specificity. Y-axis: Sensitivity.

708

709   Additional file 5 — File name: Additional\_file\_5\_Table\_S2\_exploratory\_prognostic.xlsx. File

710   format: .xlsx. Title of data: Exploratory prognostic analysis (Table S2). Description of data:

711   Cluster-based recurrence rates and logistic regression analysis of PRD and post-ablation

712   recurrence (multivariable model with 7 predefined variables).

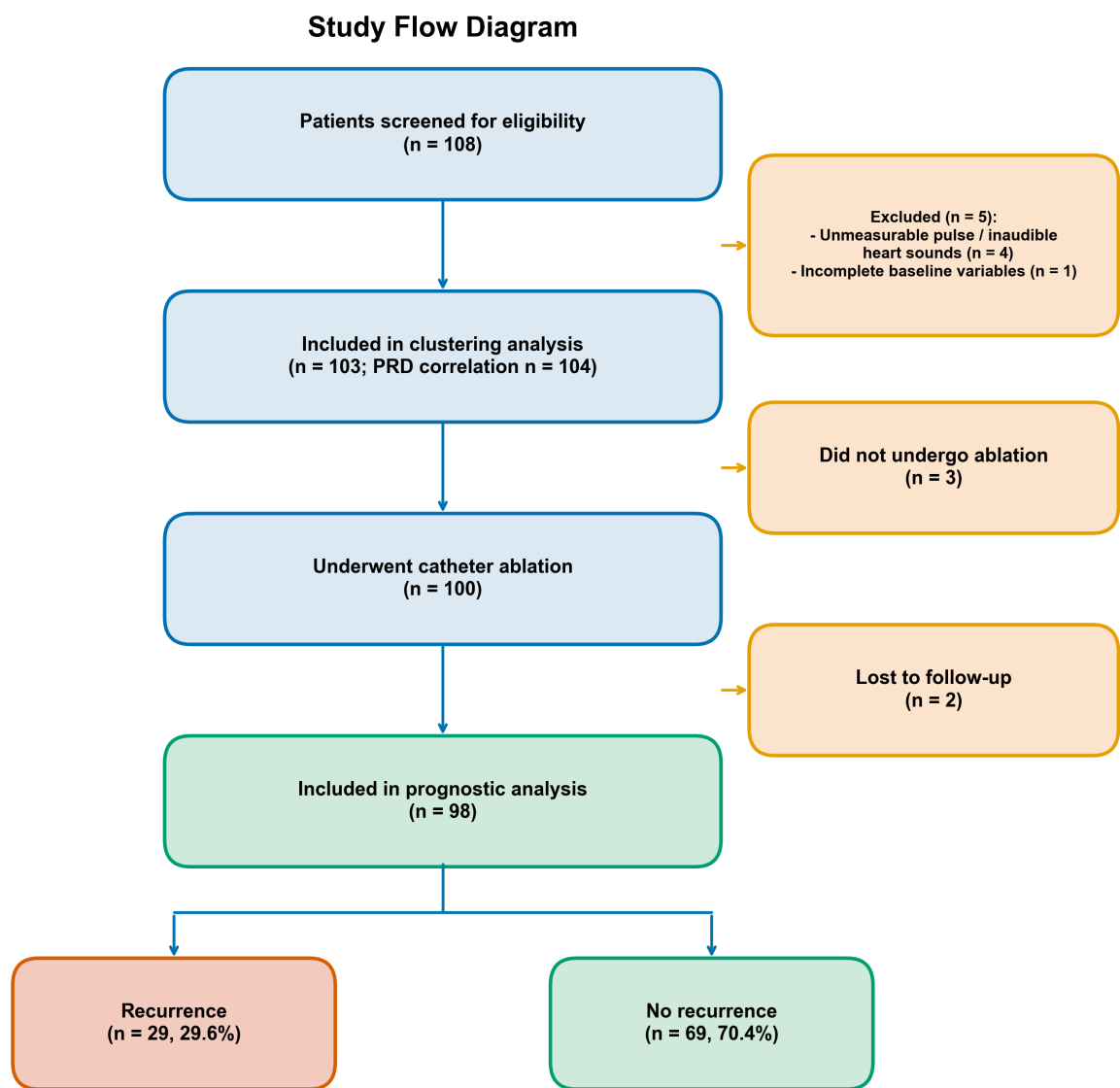
713

714   Additional file 6 — File name: Additional\_file\_6\_STROBE\_checklist.docx. File format: .docx. Title

715   of data: STROBE checklist. Description of data: Completed STROBE checklist for observational

716   studies, with the location of each item in this manuscript.

# Figures



PRD = pulse rate deficit = |auscultated heart rate – palpable pulse rate|  
PRD correlation analyses included 104 patients. One patient with valid PRD data had incomplete baseline variables and was included in the correlation analyses but excluded from clustering.

Figure 1

Study flow diagram. Flow of patients from screening ( $n=108$ ) to clustering analysis ( $n=103$ ) to prognostic analysis ( $n=98$ ). A total of 104 patients had valid PRD data and were included in the PRD correlation

analyses; one patient with valid PRD data had incomplete baseline variables and was included in the correlation analyses but excluded from the clustering analysis.

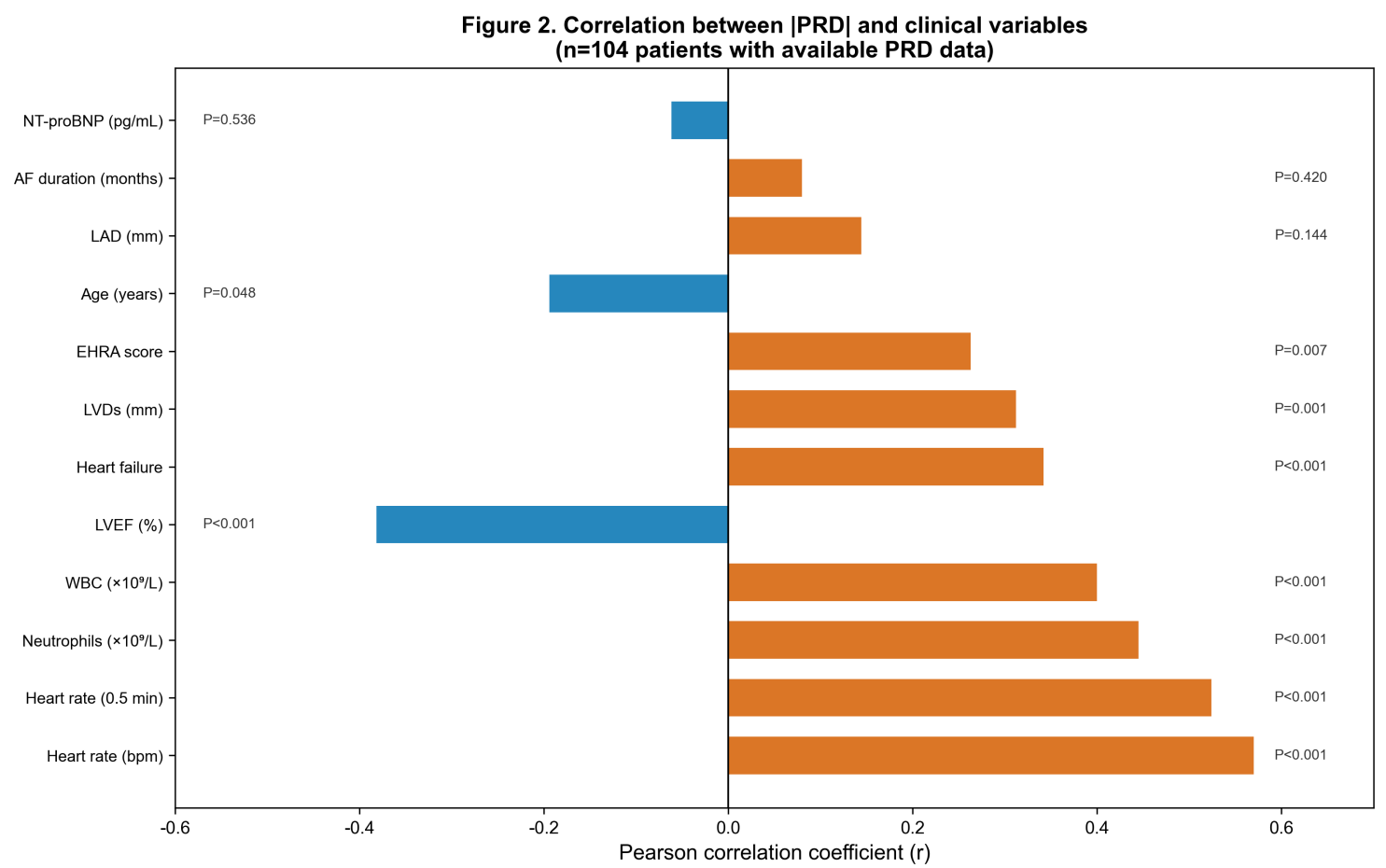


**Figure 2**

Correlation between PRD and key clinical variables. Bar plot showing Pearson correlation coefficients (r) between PRD and clinical variables, sorted by |r|. Red bars indicate positive correlation, blue bars indicate negative correlation. \*\* $P<0.01$ , \*\*\* $P<0.001$ .

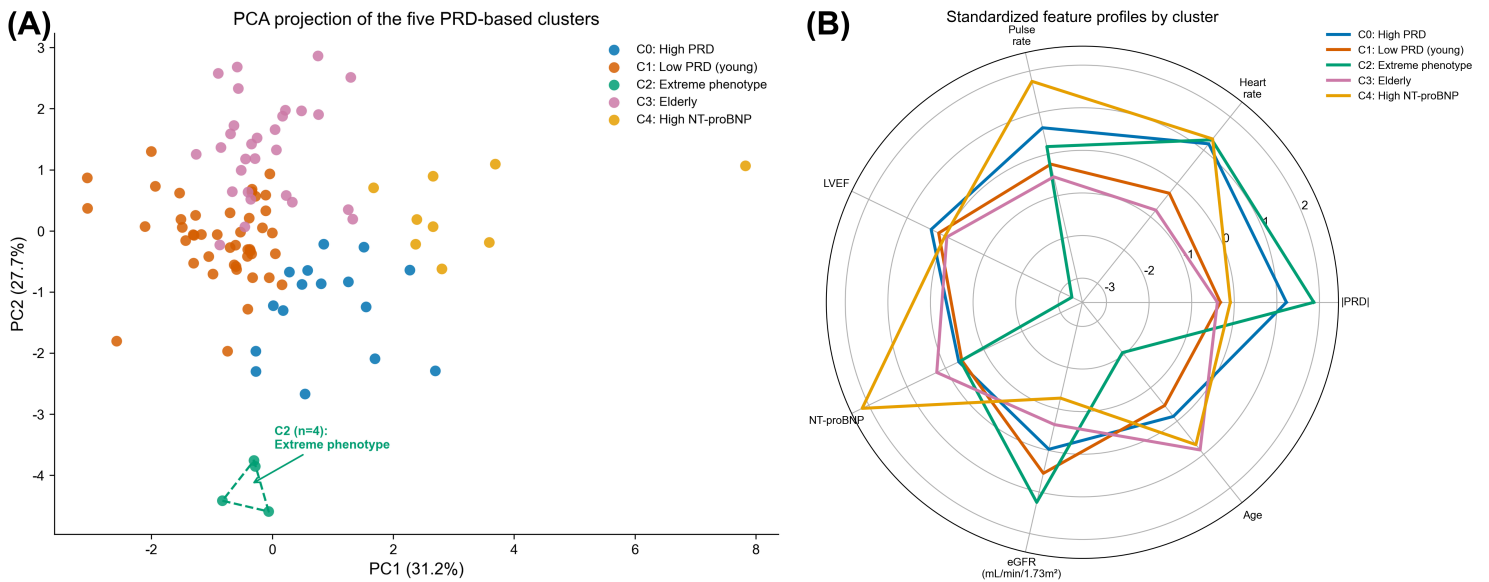


**Figure 3**

Cluster visualization. (A) Principal component analysis (PCA) scatter plot showing the distribution of 103 patients across 5 clusters in the first two principal components. Each color represents a different cluster. C2 (extreme phenotype) shows the greatest separation from other clusters. PC1 explained 31.2% and PC2 explained 27.7% of the variance; cumulative 58.9%. (B) Radar chart showing standardized feature profiles (z-scores) for each cluster across the 7 clustering variables. C2 exhibits the most deviant pattern, characterized by markedly elevated PRD, lowest LVEF, and elevated heart rate.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Table1baselinecharacteristicsv15.docx](#)
- [Table2clustercomparisonv15.docx](#)
- [Additionalfile1TableS1PRDcorrelation.xlsx](#)
- [Additionalfile5TableS2exploratoryprognostic.xlsx](#)
- [Additionalfile2FigureS1clusteringstability.png](#)
- [Additionalfile3FigureS2PRDreclassification.png](#)
- [Additionalfile4FigureS3ROCcurve.png](#)
- [Additionalfile6STROBEchecklist.docx](#)