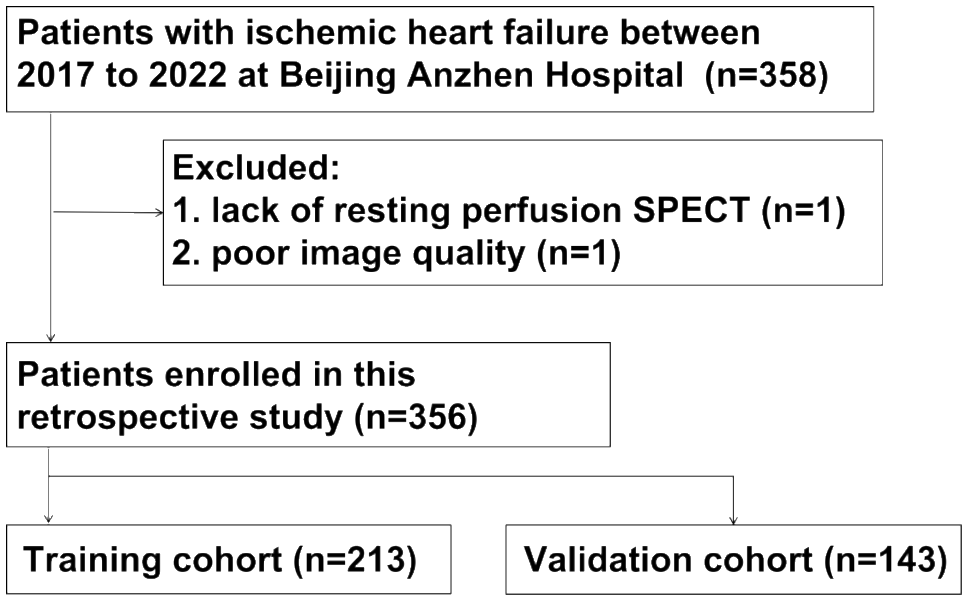
The Journal of
Cardiovascular Aging

Supplementary Material

Deep learning-based prognostic prediction in ischemic heart failure using SPECT myocardial perfusion imaging

Study Population and Enrollment

As illustrated in Supplementary Figure 1, a total of 358 patients diagnosed with ischemic heart failure (IHF) who underwent resting-state SPECT myocardial perfusion imaging (MPI) at Beijing Anzhen Hospital between 2017 and 2022 were initially screened.



Supplementary Figure 1. Flowchart of study population selection and experimental design. A total of 358 patients with ischemic heart failure (IHF) who underwent resting-state SPECT MPI at Beijing Anzhen Hospital between 2017 and 2022 were screened. Two patients were excluded: one because of lack of resting perfusion SPECT data and one because of poor image quality. The final cohort included 356 patients and was divided into a training set (n=213) and a validation set (n=143).

Inclusion Criteria: Patients were considered eligible for enrollment if they met the following criteria:

1. Confirmed Diagnosis: Diagnosed with ischemic heart failure (IHF) based on established clinical guidelines.
2. Standardized Imaging: Underwent standardized resting-state SPECT MPI using the specified protocol.
3. Data Completeness: Had complete follow-up data available for the primary endpoint analysis.

The exclusion criteria were as follows:

1. Incomplete Imaging Data: Patients lacking complete resting SPECT MPI projection data or reconstruction files.

2. Poor Image Quality: Images with significant motion artifacts, attenuation artifacts that could not be corrected, or extra-cardiac activity interfering with the left ventricular myocardium.

After applying these criteria, 2 patients were excluded: 1 because of lack of resting perfusion SPECT data and 1 because of poor image quality. The final study cohort consisted of 356 patients. This cohort was randomly divided into a training set (n=213, 60%) and a validation set (n=143, 40%) using a stratified sampling strategy to ensure a balanced distribution of event rates across both sets.

Supplementary Table 1. Additional clinical characteristics of patients in the training cohort and validation cohort

Characteristics		Training Cohort (n=213)	Validation Cohort (n=143)	P value
Smoking History	0	30 (14.1%)	26 (18.2%)	0.377
	1	29 (13.6%)	14 (9.8%)	
	unknown	154 (72.3%)	103 (72.0%)	
Drinking History	0	44 (20.7%)	34 (23.8%)	0.459
	1	15 (7.0%)	6 (4.2%)	
	unknown	154 (72.3%)	103 (72.0%)	
Family History of CAD	0	55 (25.8%)	33 (23.1%)	0.249
	1	4 (1.9%)	7 (4.9%)	
	unknown	154 (72.3%)	103 (72.0%)	
Cerebrovascular Disease	0	49 (23.0%)	35 (24.5%)	0.831

Characteristics		Training Cohort (n=213)	Validation Cohort (n=143)	P value
	1	10 (4.7%)	5 (3.5%)	
	unknown	154 (72.3%)	103 (72.0%)	
Hypertension	0	31 (14.6%)	23 (16.1%)	0.887
	1	28 (13.1%)	17 (11.9%)	
	unknown	154 (72.3%)	103 (72.0%)	
Hyperlipidemia	0	34 (16.0%)	31 (21.7%)	0.123
	1	25 (11.7%)	9 (6.3%)	
	unknown	154 (72.3%)	103 (72.0%)	
Renal Insufficiency	0	53 (24.9%)	37 (25.9%)	0.969
	1	5 (2.3%)	3 (2.1%)	
	unknown	155 (72.8%)	103 (72.0%)	
Ventricular aneurysm	0	40 (18.8%)	29 (20.3%)	0.920
	1	18 (8.5%)	11 (7.7%)	
	unknown	155 (72.8%)	103 (72.0%)	
history of myocardial infarction	0	20 (9.4%)	11 (7.7%)	0.756
	1	38 (17.8%)	29 (20.3%)	
	unknown	155 (72.8%)	103 (72.0%)	
prior PCI	0	42 (19.7%)	22 (15.4%)	0.254
	1	17 (8.0%)	18 (12.6%)	
	unknown	154 (72.3%)	103 (72.0%)	
prior CABG	0	57 (26.8%)	39 (27.3%)	0.967
	1	2 (0.9%)	1 (0.7%)	
	unknown	154 (72.3%)	103 (72.0%)	
prior CRT/ICD therapy for heart failure	0	58 (27.2%)	39 (27.3%)	0.960
	1	1 (0.5%)	1 (0.7%)	
	unknown	154 (72.3%)	103 (72.0%)	
prior revascularization therapy	0	39 (18.3%)	21 (14.7%)	0.333
	1	19 (8.9%)	19 (13.3%)	
	unknown	155 (72.8%)	103 (72.0%)	
BW		119.38 ± 41.61	117.70 ± 42.44	0.715
	Missing	9	5	
SD		29.95 (23.20-39.25)	30.55 (22.47-40.08)	0.998
	Missing	9	5	
Entropy		66.00 (60.00-72.00)	67.00 (60.00-72.00)	0.783
	Missing	52	24	

CAD, coronary artery disease; PCI, percutaneous coronary intervention; CABG, coronary artery bypass grafting; CRT, cardiac resynchronization therapy; ICD, implantable cardioverter-defibrillator; HM-TPD, heart muscle total perfusion deficit; BW, bandwidth; SD, standard deviation;

Image Acquisition and Preprocessing Details

SPECT Image Acquisition All patients underwent standardized resting gated SPECT myocardial perfusion imaging on a Siemens SPECT/CT system equipped with SMARTZOOM multifocal astigmatic collimators and IQ-SPECT technology. A cardio-centric acquisition orbit of 104° was employed with a 128 × 128 matrix. Image reconstruction was performed using the Flash-3D algorithm (10 iterations, 3 subsets) with a 10.0 mm Gaussian filter. CT-based attenuation correction and scatter correction were applied to all datasets to ensure quantitative accuracy.

Data Preprocessing for Deep Learning To standardize inputs for the deep learning model, the following preprocessing steps were applied:

1. Resampling: Original SPECT images were resampled to an isotropic voxel size of 4.795 × 4.795 × 4.795 mm.
2. Cropping: Guided by physician-annotated myocardial masks, a volume of interest (VOI) centered on the left ventricle was cropped to fixed dimensions of 40 × 50 × 50 voxels.
3. Normalization: Voxel intensities were standardized using Z-score normalization to reduce inter-subject variability.
4. Augmentation: During the training phase, on-the-fly data augmentation was applied to prevent overfitting, including random rotations (±15°), random flips, and scaling.

Network Architecture Details

The IF-SPECT model (Figure 2) employs a multi-task architecture, which integrates a 3D U-Net backbone[1],[2] with a custom Structure-Aware (SA) module.

1. Backbone & Feature Extraction:

A 3D U-Net serves as the backbone. For an input volume X, the encoder extracts multi-scale feature maps $F = \{f_1, f_2, \ldots, f_n\}$ from different levels (channels: 256, 512, 1024). Simultaneously, the decoder generates a segmentation logit map L_{seg}, from which a soft probability mask M is derived via a sigmoid function: $M = \sigma(L_{seg})$

2. Structure-Aware Attention (SAA) Module:
The SAA module fuses the encoder features with structural guidance from the segmentation mask. It comprises three key steps:

Mask-Guided Gating: Each feature map f_i is projected to a unified channel dimension ($C_{proj} = 128$) and modulated by the down-sampled mask M_i: $\hat{f}_i = Proj(f_i) \cdot (1 + \alpha \cdot M_i)$, where $\alpha = 2.0$ amplifies features in the myocardial region.

Multi-Scale Fusion & Attention: The gated features are up-sampled to a common resolution, concatenated, and fused via a $1 \times 1 \times 1$ convolution to generate F_{fused}. This feature map undergoes sequential refinement via Channel Attention (SEGate) and Spatial Attention:

a. Channel Attention (SEGate)[3]: Recalibrates channel-wise feature responses using Global Average Pooling (GAP) and a multi-layer perceptron (MLP): $F_{se} = F_{fused} \cdot \sigma(MLP(GAP(F_{fused})))$

b. Spatial Attention[4]: Highlights informative regions by aggregating channel information via Max-Pooling ($P_{\max}$) and Average-Pooling (P_{avg}), followed by a $7 \times 7 \times 7$ convolution: $M_s = \sigma(Conv^{7\times7\times7}([P_{avg}(F_{se}); P_{\max}(F_{se})]))$, $F_{refined} = F_{se} \cdot M_s$. These mechanisms effectively suppress noise and highlight salient myocardial regions.

Boundary Enhancement: To explicitly capture shape features, a fixed Laplace operator $\mathcal{L}$ (initialized with a $3 \times 3 \times 3$ edge detection kernel) is applied to the mask M to extract the myocardial boundary map M_{edge}. This is injected into the final features: $F_{final} = F_{refined} \cdot (1 + \beta \cdot \sigma(\mathcal{L}(M)))$, where $\beta = 0.5$. This ensures the prognostic features are sensitive to the myocardial geometry.

Finally, F_{final} is globally pooled and passed to a Cox Head (MLP) to predict the risk score.

Training Strategy and Hyperparameters

Two-Stage Training Strategy: To ensure robust anatomical feature learning and prevent overfitting on the small survival dataset, we employed a two-stage strategy.

Stage 1: Segmentation Pre-training (nnU-Net)

1. The 3D U-Net backbone was first trained purely for the segmentation task using the nnU-Net model. The model minimized the sum of Dice Loss and Cross-Entropy Loss to accurately delineate the Left Ventricle (LV). This stage initializes the encoder with weights capable of recognizing cardiac structures.

Stage 2: Multi-Task Fine-tuning (End-to-End)

The backbone weights from Stage 1 were loaded into the model. Unlike standard transfer learning, we did not freeze the U-Net encoder. The entire network (Encoder, SA module, and Cox Head) was fine-tuned end-to-end. To prevent catastrophic forgetting of anatomical features while learning prognosis, the model was optimized using a joint loss function consisting of both the survival loss and the segmentation loss: $L_{total} = L_{Cox} + \gamma L_{Seg}$, where L_{Cox} is the Cox negative log-likelihood loss for the survival branch, and L_{Seg} (Dice loss) maintains the segmentation quality.

Hyperparameters (Stage 2):

1. Optimizer: Adam optimizer
2. Learning Rate: 1×10^{-4}
3. Batch Size: 16
4. Epochs: 100
5. Structure Fusion Parameters: $\alpha = 2.0, \beta = 0.5$
6. Regularization: Dropout (0.2) in the Cox Head.

Supplementary Table 2. Comparison of time-dependent Receiver Operating Characteristic (ROC) curves between the IF-SPECT model and other models in the training and validation cohorts

Model		Training Cohort (n=213)	Validation Cohort (n=143)
Model-1	1-year AUC	0.727 (0.564–0.890)	0.704 (0.556–0.852)
	3-year AUC	0.813 (0.720–0.905)	0.771 (0.663–0.879)
Model-2	1-year AUC	0.775 (0.687–0.863)	0.734 (0.616–0.852)
	3-year AUC	0.754 (0.662–0.846)	0.672 (0.545–0.798)
Model-3	1-year AUC	0.673 (0.514–0.832)	0.540 (0.307–0.772)
	3-year AUC	0.726 (0.633–0.819)	0.630 (0.501–0.759)
Model-4	1-year AUC	0.623 (0.461–0.786)	0.538 (0.333–0.743)
	3-year AUC	0.712 (0.599–0.825)	0.567 (0.425–0.709)
Model-5	1-year AUC	0.740 (0.597–0.884)	0.609 (0.365–0.854)
	3-year AUC	0.831 (0.752–0.909)	0.702 (0.572–0.833)
IF-SPECT	1-year AUC	0.827 (0.734–0.920)	0.828 (0.756–0.900)
	3-year AUC	0.848 (0.778–0.918)	0.778 (0.667–0.889)

Model-1: Image-only deep learning model; Model-2: Clinical Model; Model-3: RE Score; Model-4: Radiomics Model; Model-5: Multitask Model + RE Score

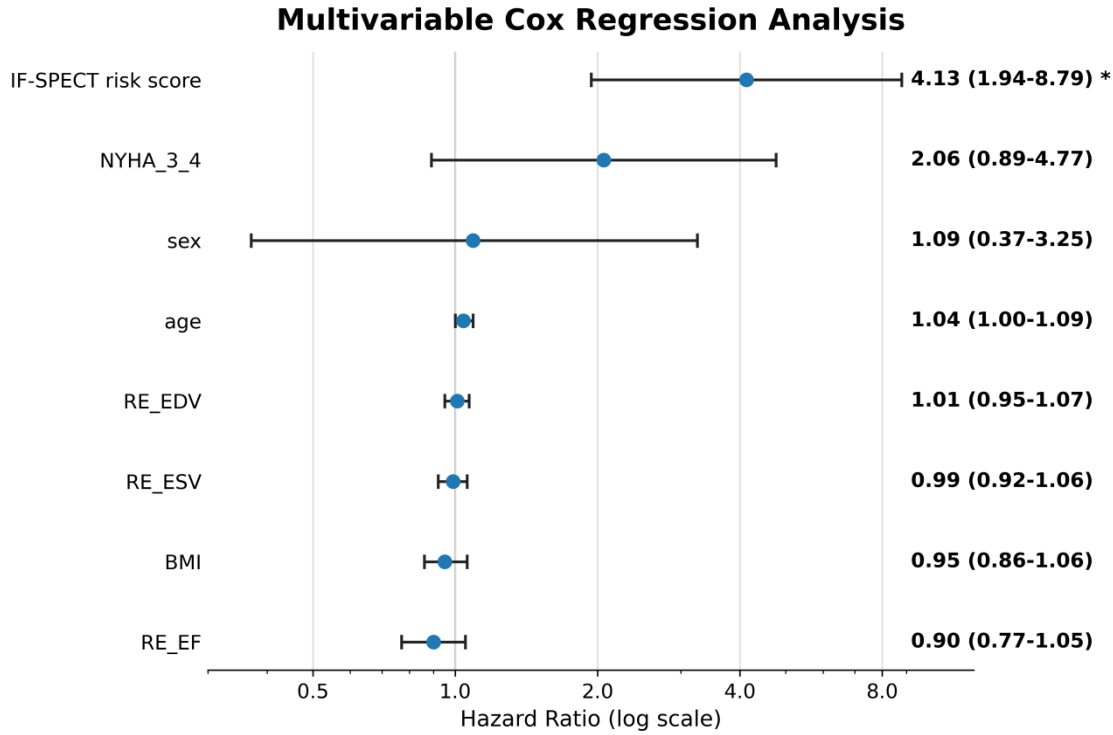
Multivariable Cox Regression Analysis

Supplementary Table 3. Multivariable Cox Regression Analysis for All-Cause Mortality

Variable	Hazard Ratio (95% CI)	P value
IF-SPECT risk score	4.13 (1.94 - 8.79)	< 0.001
NYHA_3_4	2.06 (0.89 - 4.77)	0.091
Sex	1.09 (0.37 - 3.25)	0.875
Age	1.04 (1.00 - 1.09)	0.048
RE_EDV (End-Diastolic Volume)	1.01 (0.95 - 1.07)	0.762
RE_ESV (End-Systolic Volume)	0.99 (0.92 - 1.06)	0.814

BMI	0.95 (0.86 - 1.06)	0.385
RE_EF (Ejection Fraction)	0.90 (0.77 - 1.05)	0.183

The multivariable analysis confirms that the predictive value of IF-SPECT provides incremental value beyond standard volumetric indices (EDV, ESV, EF) and functional status (NYHA).



Supplementary Figure 2. Forest plot displaying the hazard ratios (HRs) and 95% confidence intervals (CIs) for all-cause mortality. The IF-SPECT risk score remained a significant independent predictor (HR 4.13, $P < 0.001$) after adjustment for NYHA_3_4, sex, age, BMI, and standard quantitative perfusion parameters (RE_EDV, RE_ESV, RE_EF).

REFERENCES

[1] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds., Cham: Springer International Publishing, 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.

[2] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "nnU-Net:

a self-configuring method for deep learning-based biomedical image segmentation," *Nat Methods*, vol. 18, no. 2, pp. 203–211, Feb. 2021, doi: 10.1038/s41592-020-01008-z.

[3] J. Hu, L. Shen, and G. Sun, "Squeeze-and-Excitation Networks".

[4] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional Block Attention Module," in *Computer Vision – ECCV 2018*, vol. 11211, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds., in Lecture Notes in Computer Science, vol. 11211. , Cham: Springer International Publishing, 2018, pp. 3–19. doi: 10.1007/978-3-030-01234-2_1.