

Physiology-informed deep learning for early identification of sepsis-induced cardiomyopathy from echocardiographic cine loops: a multicenter development and external validation study

Yuhan Zhao

The First Affiliated Hospital with Nanjing Medical University

Qinxue Wang

The First Affiliated Hospital with Nanjing Medical University

Jiawei Wang

Affiliated Hospital of Xuzhou Medical University

Yi Han

yihan@hrbmu.edu.cn

Nanjing Medical University

Research Article

Keywords: Sepsis-induced cardiomyopathy, Echocardiography, Deep learning, Physiological modeling, External validation

Posted Date: October 30th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11318108/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

**Physiology-informed deep learning for early identification of sepsis-induced
cardiomyopathy from echocardiographic cine loops: a multicenter development
and external validation study**

Yuhan Zhao¹, Qinxue Wang^{2,3}, Jiawei Wang⁴, Yi Han^{5*}

1. The First Affiliated Hospital with Nanjing Medical University, Nanjing, Jiangsu, China.
2. Department of Geriatric Intensive Care Unit, the First Affiliated Hospital with Nanjing Medical University, Nanjing, Jiangsu, China.
3. Department of Critical Care Center, the First Affiliated Hospital with Nanjing Medical University, Nanjing, Jiangsu, China.
4. Department of Intensive Care Medicine, Affiliated Hospital of Xuzhou Medical University, Xuzhou, Jiangsu, China.
5. Nanjing Medical University, Nanjing, Jiangsu, China; Department of Critical Care Medicine, the Second Affiliated Hospital of Harbin Medical University, Harbin, Heilongjiang, China.

* Correspondence:

Yi Han

Department of Critical Care Medicine, the Second Affiliated Hospital, Harbin Medical University, Harbin, Heilongjiang, China

E-mail: yihan@hrbmu.edu.cn.

Abstract

Background

Sepsis-induced cardiomyopathy (SICM) is heterogeneous and may involve abnormalities in cardiac contraction, relaxation, and beat-to-beat motion. Conventional echocardiographic measurements and cardiac troponin T (cTnT) may not fully capture dynamic changes in myocardial function during sepsis. We developed a physiology-informed deep-learning framework for early identification of SICM from apical four-chamber (A4C) echocardiographic cine loops.

Methods

Adult patients with sepsis were retrospectively identified from two centers. The first eligible A4C cine loop obtained within 24 h after sepsis onset was used for model development. SICM was defined by cardiac abnormalities observed during the septic episode. Cine loops were divided into physiologically aligned end-diastole-to-end-diastole cardiac cycles. The model combined EchoJEPA-based visual representation learning with phase-aware spatial and temporal modeling, LV area-derived physiological features, and multi-cycle aggregation. Internal validation was performed using patient-level grouped fivefold cross-validation. The final model was trained in the full development cohort and independently evaluated in MIMIC-IV-ECHO using a conventional cTnT-based SICM phenotype. Model performance was compared with direct EchoJEPA classification.

Results

48

49 The development cohort included 530 patients, including 305 with SICM and 225 sepsis-
50 only controls. In patient-level grouped fivefold cross-validation, the physiology-informed
51 model achieved an out-of-fold AUROC of 0.933 (95% CI, 0.910–0.954). The external
52 cohort included 93 patients, including 39 with SICM and 54 sepsis-only controls. The
53 model achieved an AUROC of 0.929 (95% CI, 0.864–0.980). The direct EchoJEPA
54 baseline achieved an external AUROC of 0.735 (95% CI, 0.632–0.833). The paired
55 difference in AUROC was 0.194 (95% CI, 0.095–0.294; DeLong test, $P < 0.001$). Results
56 remained consistent in sensitivity analyses using alternative cTnT definitions.

57

58 Conclusions

59

60 A physiology-informed deep-learning framework enabled early identification of SICM
61 from routinely acquired A4C cine loops and maintained discrimination in both internal
62 cross-validation and an independent external cohort. Incorporating cardiac physiology
63 into EchoJEPA improved discrimination compared with direct foundation-model
64 classification. These findings support further prospective evaluation of physiology-
65 informed echocardiographic video analysis for early identification of SICM.

66

67 **Keywords:** Sepsis-induced cardiomyopathy; Echocardiography; Deep learning;
68 Physiological modeling; External validation.

69

70 Background

71

Sepsis-induced cardiomyopathy (SICM) is common in critically ill patients with sepsis, with a reported prevalence ranging from 10% to 70% [1, 2]. SICM has been associated with poor clinical outcomes, including significantly increased short-term mortality [1, 3]. However, its diagnosis remains difficult. Cardiac involvement during sepsis is heterogeneous and may change over time. It may present as left ventricular (LV) systolic dysfunction, diastolic dysfunction, right ventricular dysfunction, or a hyperdynamic state. There is currently no consensus on the diagnostic criteria for SICM [4]. Early recognition is important because cardiac function can help guide hemodynamic management during sepsis, including fluid, vasoactive, and inotropic therapy [5].

Evaluation of suspected SICM requires exclusion of alternative cardiac causes. Cardiac biomarkers and echocardiography are used to assess myocardial injury and cardiac function. Cardiac troponin T (cTnT) reflects myocardial injury but is not specific for SICM [6]. Echocardiography provides direct assessment of cardiac structure and function. However, conventional measurements such as LVEF, TAPSE, and diastolic indices describe only selected aspects of cardiac function. They are also affected by preload, afterload, fluid resuscitation, and vasoactive therapy [7]. Echocardiographic cine loops contain additional information. They record cardiac motion throughout the entire cardiac cycle, including contraction, relaxation, regional motion, and beat-to-beat variation [8]. These dynamic changes may not be fully captured by conventional measurements.

Deep learning provides a way to analyze spatiotemporal information directly from echocardiographic videos [9]. Video-based models have also been applied to disease-level echocardiographic detection [10]. Recent foundation models have further improved representation learning from large-scale echocardiographic data [11, 12]. More recent

multimodal approaches have extended this toward comprehensive echocardiographic interpretation [13]. However, most video-based approaches treat cine loops mainly as image sequences. Cardiac phase is not always explicitly standardized before feature extraction. Basic physiological information on ventricular contraction and filling may also need to be learned indirectly from visual features. These limitations may be particularly relevant to SICM, in which cardiac abnormalities are dynamic and heterogeneous.

We therefore developed a physiology-informed deep-learning framework based on EchoJEPA [14] for the early identification of SICM from routinely acquired apical four-chamber cine loops. The framework incorporated cardiac-cycle standardization, phase-aware spatial and temporal modeling, LV area-derived physiological features, and multi-cycle integration. We developed and trained the model in a retrospective two-center cohort and evaluated it in an independent MIMIC-IV-ECHO cohort. We hypothesized that incorporating cardiac physiology into echocardiographic video representation learning would improve the identification of SICM compared with direct EchoJEPA-based classification.

Methods

Study design and population

This retrospective study comprised a two-center development cohort and an independent external validation cohort. For the development cohort, 2,000 patients admitted to the Geriatric Intensive Care Unit of the First Affiliated Hospital with Nanjing Medical University and the Department of Critical Care Medicine of the Second Affiliated Hospital of Harbin Medical University between January 2024 and August 2026 were

retrospectively screened. Of these, 1,532 met the Sepsis-3 criteria and underwent further eligibility assessment. Eligible patients were adults aged ≥ 18 years with an ICU length of stay >24 h and <100 days. Only the first ICU admission was included for each patient. Patients were excluded if they had acute coronary syndrome, cardiomyopathy, myocarditis, clinically significant valvular heart disease, endocarditis, pericarditis, congestive heart failure, previous cardiac surgery, cardiac arrest, or severe tachyarrhythmias. Patients without sufficient echocardiographic data for SICM classification or without an analyzable apical four-chamber (A4C) cine loop were also excluded. After application of these criteria and image-quality assessment, 530 patients remained eligible for SICM classification.

Sepsis onset was defined as the first fulfillment of the Sepsis-3 criteria [15]. For each patient, the first eligible A4C cine loop obtained within 24 h after sepsis onset was selected as the index echocardiographic examination and used as model input.

SICM status was determined at the episode level according to cardiac abnormalities observed during the septic episode. SICM was defined by the presence of at least one of the following: (1) left ventricular systolic dysfunction, defined as left ventricular ejection fraction (LVEF) $<50\%$; (2) right ventricular systolic dysfunction, defined as tricuspid annular plane systolic excursion (TAPSE) <17 mm; (3) left ventricular diastolic dysfunction, defined by at least two abnormal findings among the following four criteria: reduced e' velocity (septal $e' <7$ cm/s or lateral $e' <10$ cm/s), average $E/e' >14$, peak tricuspid regurgitation velocity >2.8 m/s, and left atrial volume index >34 mL/m²; or (4) elevated high-sensitivity cardiac troponin T (hs-cTnT) >14 ng/L. Echocardiographic thresholds were based on established chamber-quantification and diastolic-function

recommendations [16, 17]. Patients who met none of the prespecified cardiac criteria were classified as sepsis-only controls. Among the 530 patients included in model development, 305 had SICM and 225 were sepsis-only controls.

External validation was performed using MIMIC-IV-ECHO, a publicly available de-identified echocardiographic dataset [18] linked to clinical records in MIMIC-IV [19]. Adult patients meeting the Sepsis-3 criteria were identified and linked to their corresponding echocardiographic examinations. The same criteria for age, first ICU admission, ICU length of stay, and cardiac exclusions were applied as in the development cohort. Patients were also required to have sufficient cTnT data for cohort classification and an analyzable A4C cine loop. The echocardiographic examination closest to sepsis onset within a ± 24 h window was selected for each patient.

In MIMIC-IV, cTnT was measured using a conventional assay. Patients with at least one cTnT concentration >0.10 ng/mL during the septic episode were classified as SICM, whereas those whose available cTnT concentrations remained ≤ 0.10 ng/mL throughout the septic episode were classified as sepsis-only controls.

Echocardiographic preprocessing and cardiac-cycle extraction

Cardiac-cycle extraction was used to standardize cardiac phase across cine loops. Each A4C cine loop was processed using a DeepLabV3-ResNet50 model for left ventricular (LV) segmentation based on the EchoNet-Dynamic framework [8]. Frames were resized to 112×112 pixels for segmentation. Channel-wise normalization parameters were

estimated from the corresponding training data and applied unchanged to subsequent evaluation data.

Frame-wise LV cavity areas were used to construct an LV area-time trajectory for each cine loop. After smoothing, end-diastolic (ED) frames were identified as local maxima of LV area. Each interval between adjacent ED frames was retained as an individual cardiac cycle. End systole (ES) was defined as the frame with the minimum LV area within each cycle. At least one complete ED-to-ED cardiac cycle was identified in each included cine loop. Each ED-to-ED cycle was resampled to 16 frames and resized to 224×224 pixels for subsequent modeling.

LV area-based physiological representation

The LV area-time trajectory was also used to derive a physiological representation for each cardiac cycle. The LV area curve was normalized by the mean area of the two ED endpoints and interpolated to 64 equally spaced temporal positions.

Five physiological descriptors were calculated: LV fractional area change (LV-FAC), normalized systolic contraction rate, normalized diastolic filling rate, a frame-based cycle-length proxy, and relative ED endpoint mismatch. The five descriptors were standardized using feature-wise means and standard deviations estimated from the corresponding training data and applied unchanged to evaluation data. The normalized LV area curve and the scalar descriptors were encoded separately using multilayer perceptrons and combined into a 128-dimensional physiological representation.

EchoJEPa encoder and self-supervised adaptation

EchoJEPa, an echocardiographic foundation model with a ViT-Large backbone [14], was used for visual feature extraction and initialized from the pretrained vitl-vmix22m-pt220-c55 checkpoint. Each cardiac cycle contained 16 grayscale frames at 224×224 pixels. The frames were replicated across three channels and normalized using the preprocessing parameters of the pretrained EchoJEPa model.

The encoder used a 16×16 spatial patch size and a temporal tubelet size of two frames, producing eight temporal positions with a 14×14 spatial-token grid at each position. Before supervised SICM classification, the pretrained encoder underwent self-supervised adaptation using cardiac cycles from the corresponding training data.

Self-supervised adaptation combined masked latent prediction within each cardiac cycle and cross-cycle consistency between cycles from the same A4C cine loop. For masked latent prediction, 60% of the spatiotemporal tokens were masked. The online encoder predicted the masked latent representations, while an exponential-moving-average target encoder generated the corresponding targets from an independently augmented view of the same cycle. Smooth L1 loss was calculated between L2-normalized predicted and target representations.

When multiple cardiac cycles were available from the same cine loop, a different cycle was used for cross-cycle consistency. Cosine similarity was used to compare the cycle-level representations. The overall self-supervised objective was

$$\mathcal{L}_{SSL} = \mathcal{L}_{pred} + 0.25 \mathcal{L}_{cycle}.$$

222

223 Self-supervised adaptation was performed for 10 epochs. The final four Transformer
 224 blocks and the final normalization layer of EchoJEPA were trainable, together with the
 225 spatial- and temporal-attention modules.

226

227 **Phase-aware modeling and multiple-instance learning**

228

229 EchoJEPA token representations were organized into eight temporal positions, with 196
 230 spatial tokens at each position. Phase-aware spatial attention assigned weights to the
 231 spatial tokens and generated one representation for each temporal position. Phase-
 232 aware temporal attention then modeled information across the eight positions and
 233 generated a 1024-dimensional cycle-level visual representation.

234

235 For each cardiac cycle, the visual representation and the 128-dimensional physiological
 236 representation together formed the cycle representation. All available cardiac cycles
 237 from the index A4C cine loop were then analyzed using multiple-instance learning (MIL).
 238 MIL assigned an attention weight to each cycle and aggregated the cycle
 239 representations into a patient-level representation:

240

$$241 \quad \mathbf{H} = \sum_{k=1}^K \alpha_k \mathbf{h}_k,$$

242

243 where $\mathbf{h}_k$ represents cardiac cycle k and α_k its normalized attention weight. The
 244 aggregated representation was passed to a multilayer perceptron classifier to generate

the patient-level probability of SICM. Classification was performed only at the patient level.

Data augmentation and supervised training

Temporally consistent augmentation was applied to all frames within each cardiac cycle during model development. Augmentations included changes in brightness, contrast, gamma, and small spatial translations.

For the final model, the entire development cohort was used for model training. Supervised training consisted of two stages. During Stage A, the EchoJEPA backbone was frozen for 10 epochs while the downstream modules were trained. During Stage B, the final two EchoJEPA Transformer blocks and the final normalization layer were unfrozen and jointly fine-tuned for an additional 20 epochs. AdamW optimization, cosine learning-rate annealing, and binary cross-entropy with logits were used.

Internal validation by patient-level grouped fivefold cross-validation

Internal validation was performed using patient-level grouped fivefold cross-validation in the development cohort. Patients were assigned to five mutually exclusive folds with similar SICM proportions, and all cardiac cycles from the same patient were assigned to the same fold. In each iteration, four folds were used for training and the remaining fold for validation.

For each fold, self-supervised adaptation, physiological-feature normalization, and supervised training were performed using only the training folds. The model was initialized from the same pretrained EchoJEPA checkpoint and trained using the same procedures described above. Predictions from the five held-out folds were combined to generate one out-of-fold probability for each patient. The pooled out-of-fold AUROC was used to assess internal discrimination.

External validation and sensitivity analyses

The final model and preprocessing pipeline were fixed and applied to the MIMIC-IV-ECHO cohort without further model updating. External echocardiographic videos underwent the same LV segmentation, cardiac-cycle extraction, temporal and spatial resampling, physiological-feature generation, and MIL inference procedures used during model development. LV-segmentation normalization parameters were derived from the development cohort, whereas physiological normalization parameters were stored within the trained model and applied unchanged during external inference. No random augmentation was applied during external evaluation. For each patient, all cardiac cycles extracted from the selected A4C cine loop were jointly analyzed to generate a single SICM probability.

The primary external analysis defined SICM using a cTnT cutoff of >0.10 ng/mL. Two sensitivity analyses were performed using the same model predictions. First, the cTnT cutoff was increased to >0.20 ng/mL. Second, patients with cTnT concentrations between 0.08 and 0.12 ng/mL were excluded, and the primary 0.10 ng/mL cutoff was retained for classification of the remaining patients.

Model evaluation

A probability threshold of 0.5 was prespecified for binary classification. Model discrimination was assessed using the area under the receiver operating characteristic curve (AUROC). For fivefold cross-validation, predictions from the held-out fold in each iteration were pooled to calculate the patient-level out-of-fold AUROC. Fold-specific AUROCs were also calculated. Accuracy, sensitivity, specificity, and F1 score were calculated at the prespecified threshold. Calibration was assessed using a calibration curve and the Brier score. The same patient-level model probabilities were used for the primary external analysis and both sensitivity analyses; only the corresponding reference labels were changed.

Statistical analysis

Continuous variables are presented as median (interquartile range), and categorical variables as number (percentage). Continuous variables were compared using the Mann–Whitney *U* test, whereas categorical variables were compared using the chi-square test or Fisher’s exact test, as appropriate.

For analyses of LV area-derived physiological features, measurements from multiple cardiac cycles were aggregated within patient before group-level comparison. 95% confidence intervals for AUROC were estimated using bootstrap resampling. For fivefold cross-validation, the same procedure was applied to the pooled out-of-fold AUROC. The physiology-informed model and EchoJEPA baseline were compared in the external

cohort using the paired DeLong test. Net benefit across threshold probabilities was assessed using decision curve analysis. All statistical tests were two-sided, and $P < 0.05$ was considered statistically significant.

Results

Study population

A total of 2,000 patients were screened across the two participating centers. Among them, 1,532 met the Sepsis-3 criteria. After clinical and echocardiographic eligibility assessment, 530 patients had sufficient data for SICM classification. Of these, 305 were classified as SICM and 225 as sepsis-only controls (Fig. 1). The development cohort had a median age of 60 years, and 65.8% were male. Clinical and laboratory characteristics on the day of sepsis diagnosis are summarized in Table 1. Patients with SICM had slightly higher SOFA scores than sepsis-only controls (11 [8–13] vs. 10 [7–12], $P=0.02$).

Table 1 Baseline characteristics of the development cohort

Characteristic	Overall (n = 530)	Sepsis-only (n = 225)	SICM (n = 305)	<i>P</i> value
Demographics				
Age, years	60 [51, 70]	59 [52, 68]	61 [50, 71]	0.83
Male, n (%)	349 (65.8%)	139 (61.8%)	210 (68.9%)	0.09
Comorbidities				
Diabetes mellitus, n (%)	95 (17.9%)	40 (17.8%)	55 (18.0%)	0.94
Chronic kidney disease, n (%)	63 (11.9%)	30 (13.3%)	33 (10.8%)	0.38
Chronic liver disease, n (%)	45 (8.5%)	20 (8.9%)	25 (8.2%)	0.78
Severity of illness				

SOFA score	11 [8, 13]	10 [7, 12]	11 [8, 13]	0.02
APACHE II score	14 [11, 20]	14 [10, 20]	14 [11, 21]	0.17
Sites of infection				
Respiratory infection, n (%)	244 (46.0%)	95 (42.2%)	149 (48.9%)	0.13
Abdominal infection, n (%)	127 (24.0%)	61 (27.1%)	66 (21.6%)	0.15
Bloodstream infection, n (%)	145 (27.4%)	69 (30.7%)	76 (24.9%)	0.14
Other infection, n (%)	107 (20.2%)	35 (15.6%)	72 (23.6%)	0.02
Laboratory findings				
White blood cell count, $\times 10^9/L$	9.59 [6.71, 12.74]	9.75 [6.72, 13.22]	9.18 [6.70, 12.53]	0.34
Procalcitonin, ng/mL	2.10 [0.92, 5.38]	2.01 [0.76, 5.32]	2.14 [1.04, 5.65]	0.23
C-reactive protein, mg/L	77.60 [30.70, 130.00]	83.10 [31.20, 135.00]	76.50 [29.62, 126.00]	0.54
Total bilirubin, $\mu\text{mol/L}$	13.50 [9.10, 19.90]	13.25 [8.68, 19.98]	13.80 [9.25, 19.70]	0.60
Creatinine, $\mu\text{mol/L}$	82.70 [56.30, 127.35]	84.35 [54.40, 122.90]	81.85 [57.08, 128.32]	0.50

Data are presented as median [interquartile range] or n (%). For variables with missing data, analyses were based on available cases. Clinical severity scores and laboratory measurements refer to values obtained on the day of sepsis diagnosis. Infection sites were not mutually exclusive. APACHE II, Acute Physiology and Chronic Health Evaluation II; SICM, sepsis-induced cardiomyopathy; SOFA, Sequential Organ Failure Assessment.

The cardiac abnormalities contributing to SICM varied across patients. Among patients with SICM, left ventricular diastolic dysfunction was present in 183/297 (61.6%), LVEF <50% in 176/301 (58.5%), TAPSE <17 mm in 128/291 (44.0%), and hs-cTnT >14 ng/L in 127/305 (41.6%) (Table S1).

Table S1 Cardiac abnormalities among patients with SICM

Characteristic	SICM (n = 305)	Abnormal, n/N (%)
LVEF, %	41.60 [29.30,	176 (58.5%)

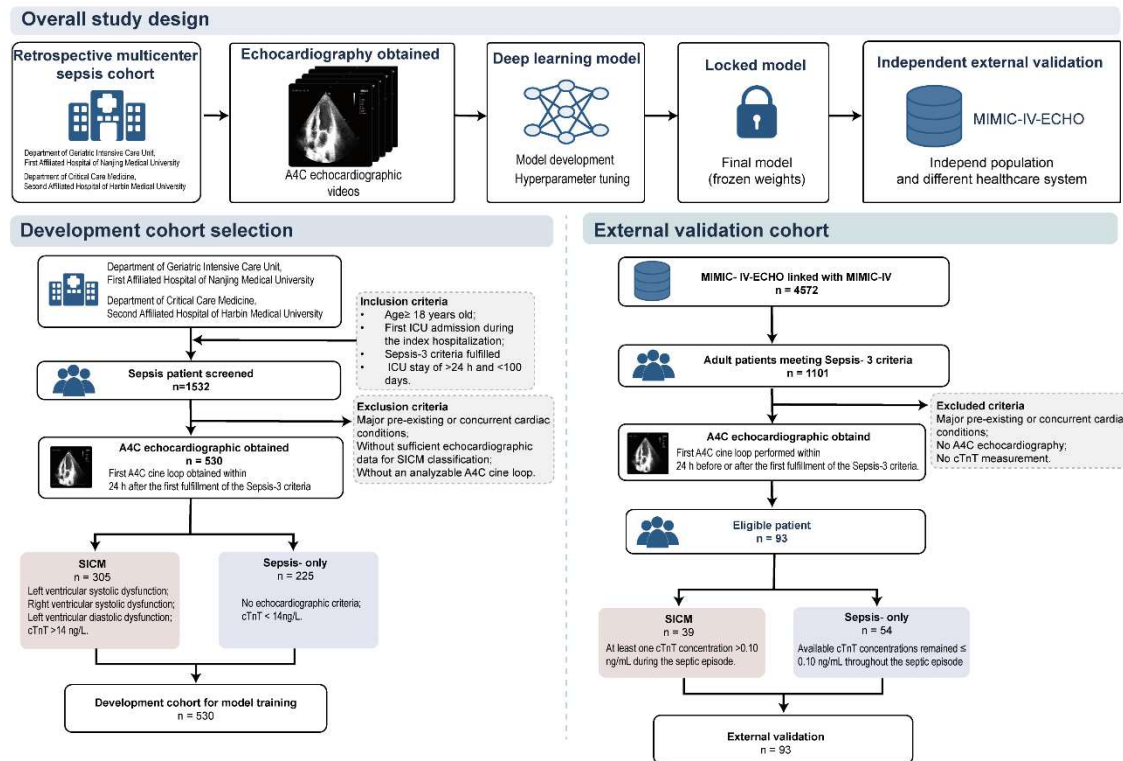
	53.50]	
TAPSE, mm	16.85 ± 3.74	128 (44.0%)
Left ventricular diastolic dysfunction, n (%)		231 (78.8%)
Septal e', cm/s	6.50 ± 1.62	168 (57.3%)
Lateral e', cm/s	8.62 ± 2.27	196 (68.1%)
Average E/e'	11.10 [8.50, 15.10]	77 (28.6%)
Peak TR velocity, m/s	3.13 ± 0.51	190 (72.2%)
LAVI, mL/m ²	41.48 ± 9.88	219 (77.7%)
hs-cTnT, ng/L	12.01 [6.39, 53.91]	129 (42.3%)

Continuous variables are presented as mean ± standard deviation or median [interquartile range], as appropriate. Abnormal values are presented as n/N (%), where N represents the number of patients with an available measurement. LVEF, left ventricular ejection fraction; TAPSE, tricuspid annular plane systolic excursion; TR, tricuspid regurgitation; LAVI, left atrial volume index

340

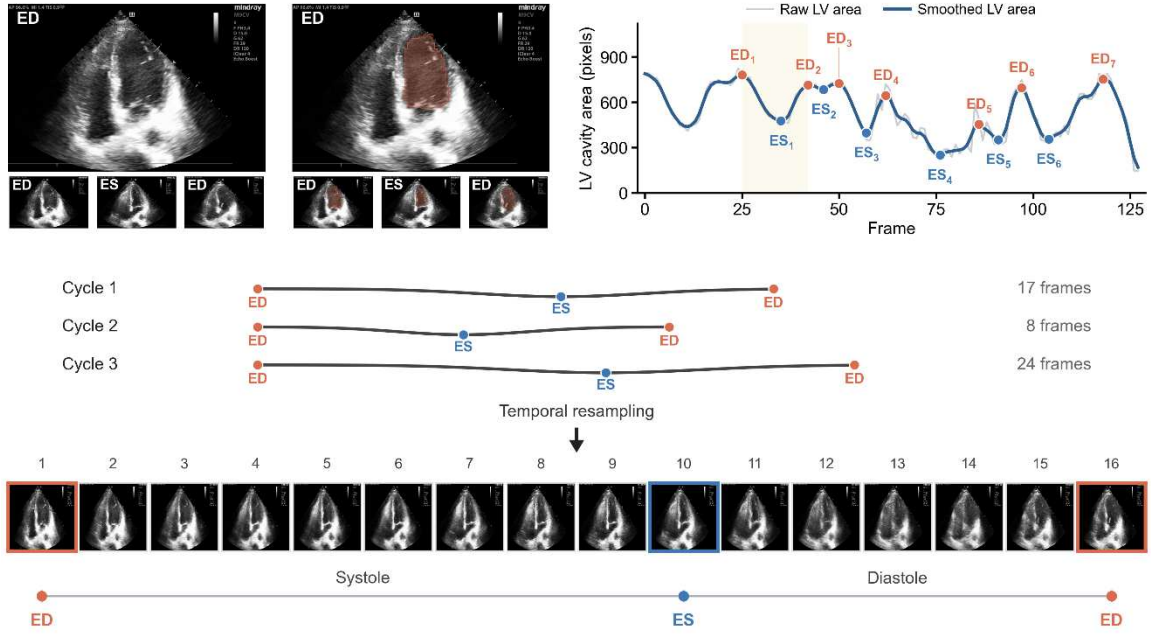
341

342 The external validation cohort included 93 eligible patients from MIMIC-IV-ECHO, of
 343 whom 39 (41.9%) met the cTnT-based SICM definition and 54 were classified as sepsis-
 344 only controls (Fig. 1).

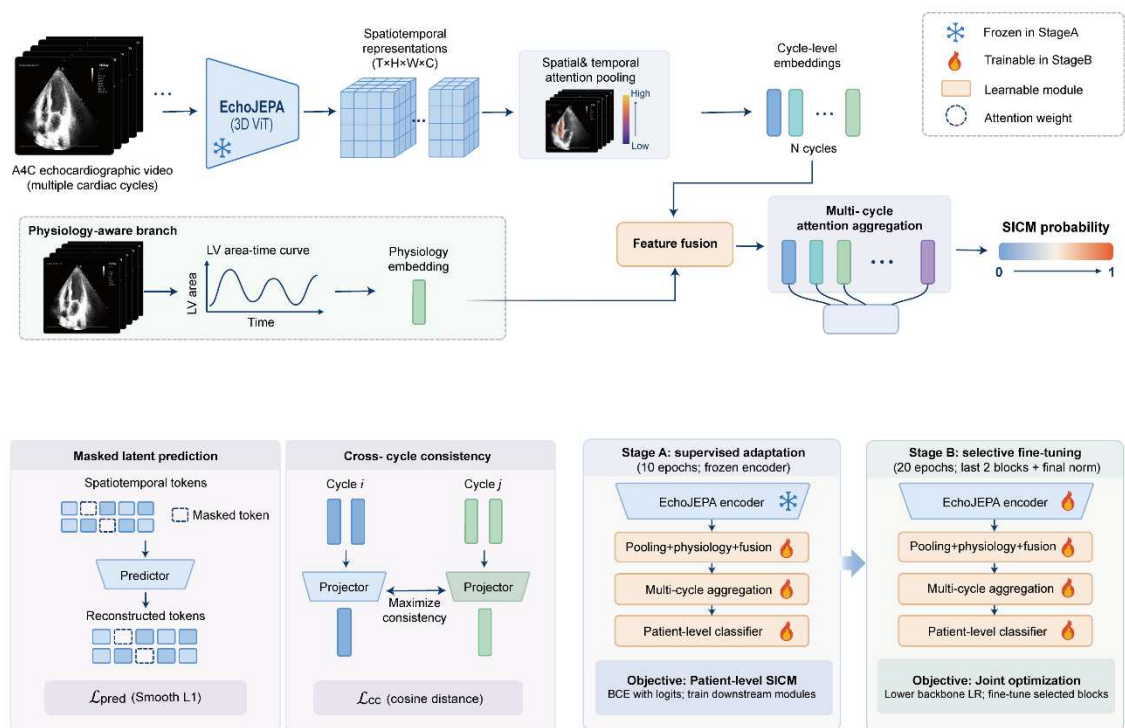


Physiology-informed cardiac-cycle modeling

Physiological cardiac-cycle alignment was introduced to account for differences in cardiac phase and cycle duration across cine loops. Automated LV segmentation generated an area-time trajectory for each A4C cine loop. ED and ES were identified from this trajectory, and adjacent ED-to-ED intervals were extracted as individual cardiac cycles. Each cycle was resampled to 16 frames, providing a common physiological sequence from ED through systole and ES to diastole and the subsequent ED (Fig. 2).



Each standardized cycle was represented using both visual and physiological information. EchoJEPA extracted high-dimensional spatiotemporal features from each standardized cycle. These representations were further adapted to cardiac-cycle dynamics using masked latent prediction within individual cycles and cross-cycle consistency between different heartbeats from the same cine loop. Phase-aware spatial and temporal attention was then applied to the EchoJEPA features to generate the cycle-level visual representation. In parallel, the LV area trajectory provided an explicit physiological representation of cardiac contraction and filling. The visual and physiological representations together formed the cycle-level representation. Multiple cardiac cycles from the same A4C cine loop were retained. MIL assigned different weights to the cycle representations and combined them into a patient-level representation for SICM classification (Fig. 3).



369

370

371 **Model performance and external validation**

372

373 The physiology-informed model achieved an apparent AUROC of 0.978 (95% CI, 0.966–
374 0.988) in the development cohort. In patient-level grouped fivefold cross-validation, the
375 pooled out-of-fold AUROC was 0.933 (95% CI, 0.910–0.954) (Table S2). Performance
376 remained high in the independent MIMIC-IV-ECHO cohort, with an AUROC of 0.929 (95%
377 CI, 0.864–0.980) (Fig. 4a). The model correctly classified 35 of 39 patients with SICM
378 and 49 of 54 sepsis-only controls. The accuracy was 90.3%, sensitivity 89.7%, specificity
379 90.7%, and F1 score 88.6% (Fig. 4e). Calibration between predicted probabilities and
380 observed outcomes was good, with a Brier score of 0.083 (Fig. 4b). Decision curve

analysis showed a higher net benefit for the physiology-informed model than for the EchoJEPA baseline across the evaluated threshold probabilities (Fig. 4c).

Table S2. Patient-level fivefold cross-validation performance in the development cohort

Fold	Held-out fold, n (SICM/sepsis- only)	AUROC (95% CI)	Sensitivity	Specificity	Brier score
1	106 (61/45)	0.939 (0.886–0.977)	0.934	0.889	0.088
2	106 (61/45)	0.927 (0.869–0.975)	0.918	0.889	0.097
3	106 (61/45)	0.920 (0.853–0.974)	0.902	0.844	0.115
4	106 (61/45)	0.932 (0.882–0.973)	0.869	0.911	0.115
5	106 (61/45)	0.970 (0.937–0.994)	0.885	0.956	0.076
Mean ± SD	—	0.938 ± 0.019	0.902 ± 0.026	0.898 ± 0.040	0.098 ± 0.017
Pooled out-of-fold	530 (305/225)	0.933 (0.910–0.954)	0.902	0.898	0.098

Data are presented for each held-out fold, as mean ± SD across the five folds, and for the pooled out-of-fold predictions. Each patient contributed to the pooled analysis only once, when assigned to the held-out fold. AUROC 95% CIs were estimated using 2,000 bootstrap resamples. Sensitivity and specificity were calculated using a probability threshold of 0.5. AUROC, area under the receiver operating characteristic curve; CI, confidence interval; OOF, out-of-fold; SD, standard deviation.

The advantage over direct EchoJEPA classification was also observed in both cohorts. The EchoJEPA baseline achieved an apparent AUROC of 0.816 (95% CI, 0.782–0.849) in the development cohort and an AUROC of 0.735 (95% CI, 0.632–0.833) in the external cohort. In the external cohort, the paired difference in AUROC was 0.194 (95% CI, 0.095–0.294; DeLong test, $P < 0.001$) (Fig. 4d,f).

The findings were maintained when alternative cTnT definitions were used. With a cTnT cutoff of >0.20 ng/mL, 22 of 93 patients were classified as SICM and the AUROC was 0.808 (95% CI, 0.715–0.889). After excluding patients with cTnT concentrations of 0.08–0.12 ng/mL, 33 of 84 patients were classified as SICM and the AUROC was 0.923 (95% CI, 0.853–0.979) (Table S3).

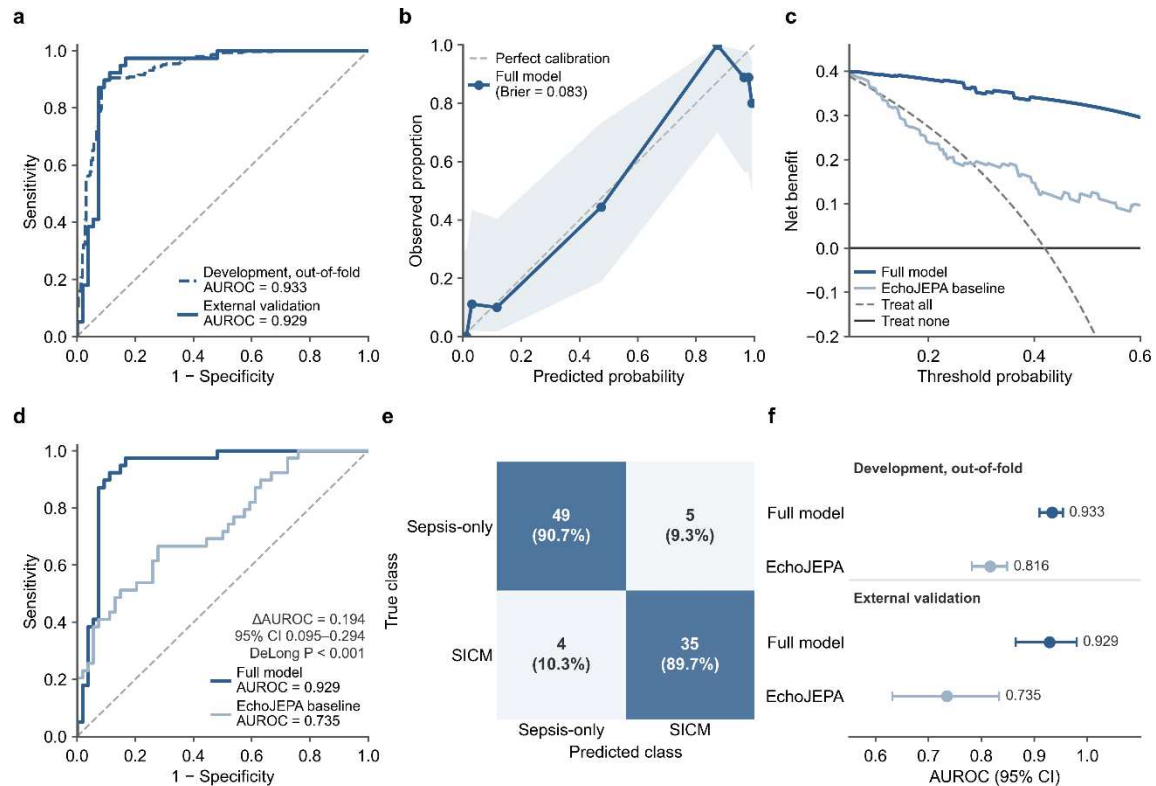


Table S3 Sensitivity analyses of alternative cTnT definitions

Analysis	cTnT definition	Additional restriction	SICM, n/N (%)	AUROC (95% CI)
Primary analysis	>0.10 ng/mL	None	39/93 (41.9%)	0.929 (0.864–0.980)
Sensitivity analysis 1	>0.20 ng/mL	None	22/93 (23.7%)	0.808 (0.715–0.889)
Sensitivity analysis 2	>0.10 ng/mL	Excluded cTnT 0.08–0.12 ng/mL	33/84 (39.3%)	0.923 (0.853–0.979)

All analyses used the same model predictions; only the reference classification changed. SICM status was based on conventional cTnT measured during the septic episode. 95% CIs were estimated by bootstrap resampling. AUROC, area under the receiver operating characteristic curve; CI, confidence interval; cTnT, cardiac troponin T; SICM, sepsis-induced cardiomyopathy.

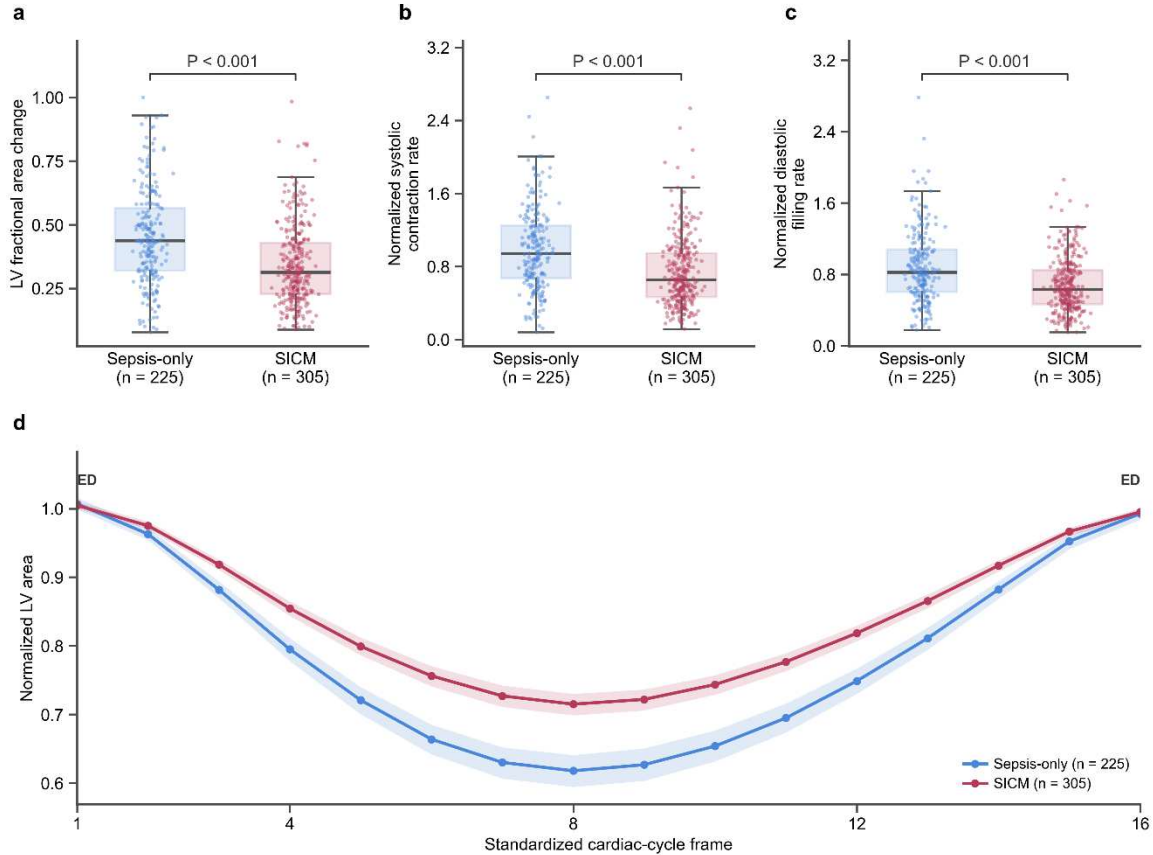
399

400

401 **Physiological and attention-based interpretation**

402

403 Physiological and attention analyses were used to characterize the cardiac-motion
404 information represented by the model. The aligned cardiac cycles showed clear
405 differences in cardiac mechanics between patients with SICM and sepsis-only controls.
406 LV-FAC, systolic contraction rate, and diastolic filling rate were all lower in SICM (all
407 $P < 0.001$; Fig. 5a–c). The normalized LV area trajectory also showed less systolic area
408 reduction and a different pattern during diastolic filling in SICM (Fig. 5d). However, the
409 distributions of individual physiological measures overlapped substantially between
410 groups.



411

412

413 Attention analysis further showed how this information was distributed within the model.

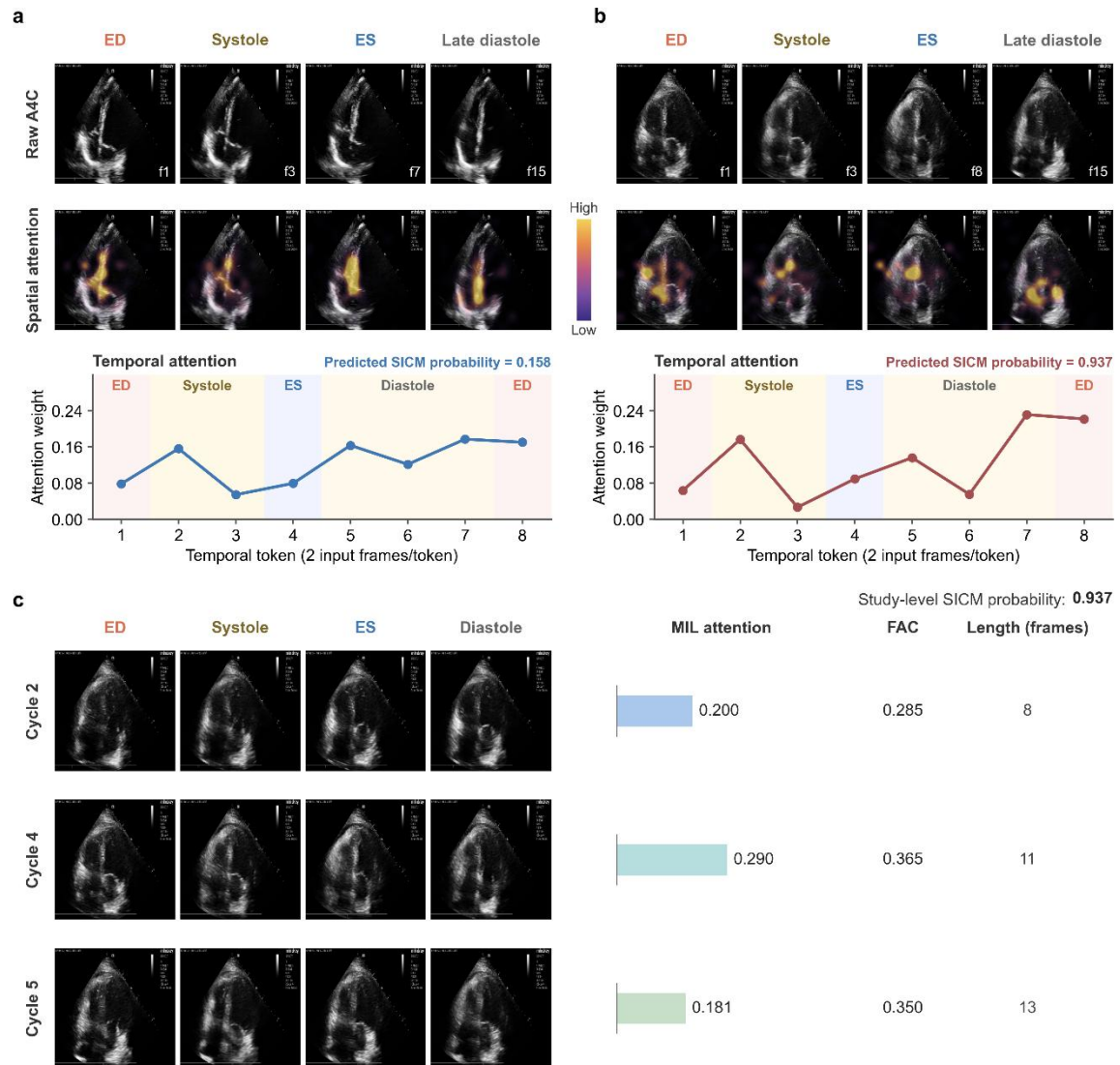
414 Spatial attention varied across cardiac phases, and temporal attention was distributed

415 differently within individual cardiac cycles (Fig. 6a,b). MIL also assigned different weights

416 to individual heartbeats within the same examination (Fig. 6c). These findings showed

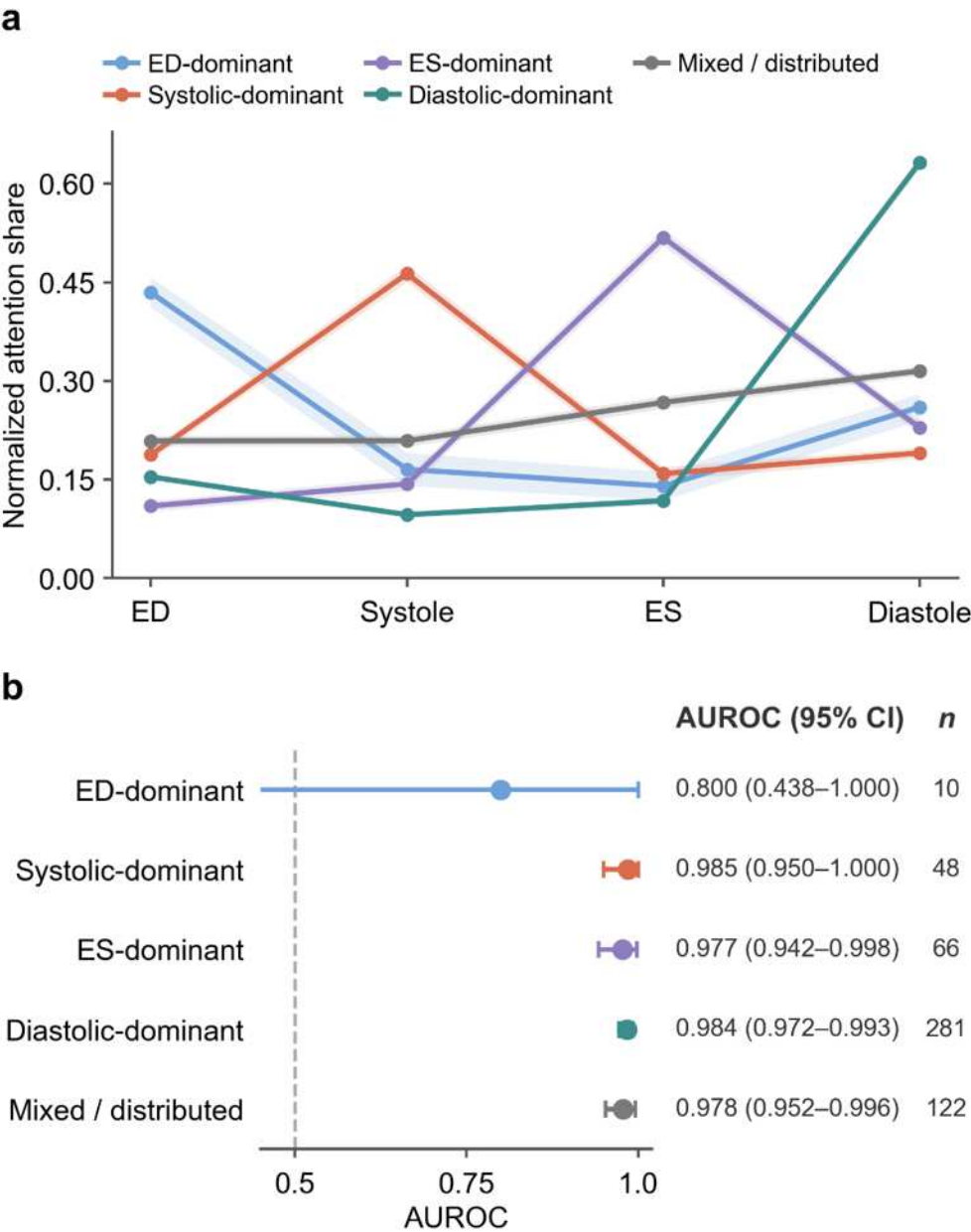
417 that the patient-level prediction was based on information from different spatial regions,

418 cardiac phases, and cardiac cycles.



In further analysis, we found that temporal attention was distributed differently across the cardiac cycle among patients. Five patterns were observed: ED-dominant, systolic-dominant, ES-dominant, diastolic-dominant, and mixed/distributed attention (Fig. S1). Representative examinations from these subgroups showed distinct spatial-attention patterns across cardiac phases (Fig. S2). Discrimination remained high in the systolic-dominant, ES-dominant, diastolic-dominant, and mixed/distributed groups, with AUROCs ranging from 0.977 to 0.985. The smaller ED-dominant group had an AUROC of 0.800.

428 These findings suggest that the model can identify SICM-related motion abnormalities
429 arising from different phases of the cardiac cycle.



430

0.929 in the independent MIMIC-IV-ECHO cohort. By comparison, direct classification with the same pretrained EchoJEPa encoder achieved an external AUROC of 0.735. The improvement in model performance suggests that explicitly incorporating cardiac-cycle structure and ventricular mechanics provides information beyond pretrained visual representations alone. Physiological and attention analyses further showed that model-relevant information was distributed across cardiac phases, ventricular motion, and successive heartbeats.

The framework was designed around the way cardiac function is assessed in clinical echocardiography. Clinical interpretation is organized around the cardiac cycle, including identification of end-diastole and end-systole, assessment of ventricular contraction and filling across cardiac phases, and consideration of more than one heartbeat when beat-to-beat variation is present. Deep-learning models and echocardiographic foundation models can learn powerful representations directly from video, but they usually receive a fixed number of frames sampled from each clip, so this physiological structure must be learned implicitly [20]. In our framework, individual heartbeats were aligned to a common cardiac cycle, ventricular mechanics were represented explicitly, visual information was modeled according to cardiac phase, and multiple heartbeats were integrated at the patient level. Because the comparison model used the same pretrained EchoJEPa encoder, the observed performance difference suggests additional value from the physiology-informed modeling strategy beyond the pretrained visual representation alone.

Cardiac-cycle alignment follows directly from this principle. Cine loops differ in acquisition length, starting position, heart rate, and cardiac phase, so when frames are

sampled at fixed intervals, the same temporal position may correspond to early systole in one patient and mid-diastole in another. Using the LV area trajectory, we identified ED from local maxima in ventricular area and used consecutive ED frames to define individual cardiac cycles. Each ED-to-ED interval was then resampled to a fixed length, providing a common physiological time axis in which each position represented a corresponding stage of contraction or filling across patients. This aligned systolic contraction and diastolic filling to corresponding positions within the standardized cycle and reduced temporal misalignment caused by differences in cycle duration and video acquisition. Cycle length itself was retained as a physiological descriptor. This matters in sepsis, where abnormalities may involve systolic contraction, diastolic filling, right ventricular function, or combinations of these [21].

The physiological branch provided the model with the same kind of mechanical information that clinicians derive from ventricular size over time. For each aligned cycle, the normalized LV area trajectory and descriptors of contraction, filling, and cycle length were encoded alongside the visual features. Consistent with the expected physiological pattern, LV-FAC, systolic contraction rate, and diastolic filling rate were all lower in patients with SICM than in sepsis-only controls. Their distributions nevertheless overlapped substantially, consistent with the heterogeneous cardiac phenotypes of sepsis, and no single measure identified SICM on its own [22]. This mirrors clinical experience with conventional indices. The physiological branch was therefore not intended as a standalone classifier. It provided the model with an explicit, interpretable representation of ventricular mechanics that was combined with visual information from the full cardiac cycle.

Phase-aware modeling allowed the model to distinguish information from different parts of the cardiac cycle, while multi-cycle integration incorporated variation between heartbeats. With the cycles aligned, spatial and temporal attention could weigh information within physiologically defined phases, analogous to assessing systolic and diastolic function separately. Because cardiac motion in critically ill patients varies from beat to beat with heart rate, respiration, and ventricular filling, all aligned cycles were retained and combined rather than relying on a single, possibly unrepresentative beat [23]. Attention varied across cardiac phases, spatial regions, and heartbeats, and discrimination remained high across the phase-specific attention subgroups in the supplementary analyses. Attention weights are not direct evidence of biological mechanisms, but they indicate that the model drew on information distributed across the cardiac cycle and across beats rather than on a single phase or measurement [24].

The validation strategy supports these findings. Apparent performance was supplemented by patient-level grouped fivefold cross-validation, in which every patient was evaluated by a model that had not seen that patient during training [25]. The pooled out-of-fold AUROC was 0.933 (95% CI, 0.910–0.954), compared with an apparent AUROC of 0.978 in the full development cohort. After training on the full development cohort, the model was locked and applied to MIMIC-IV-ECHO without further updating [26]. The model achieved an external AUROC of 0.929 (95% CI, 0.864–0.980) and continued to perform well in both sensitivity analyses (AUROCs 0.808 and 0.923). These findings suggest that the model retained its ability to distinguish cases from controls beyond the development cohort, providing support for the transportability of the physiology-informed approach [27].

The clinical setting of the model is early sepsis. Decisions about fluids, vasopressors, and inotropes are made early in the course of sepsis, when loading conditions are changing and conventional measurements such as LVEF reflect load as well as intrinsic myocardial function. Troponin indicates myocardial injury but not its functional consequences. Bedside echocardiography is widely used at this stage, and because our model needs only a standard A4C cine loop, it can be applied to the first routine examination without additional imaging [28]. An early model-based signal could help identify patients who warrant closer cardiovascular assessment, repeat echocardiography, or more careful titration of hemodynamic therapy. Whether this improves management or outcomes requires prospective evaluation [29].

This study has several limitations. First, the development cohort was retrospective and included two centers, and examinations were obtained during routine care rather than under a standardized imaging protocol; broader multicenter evaluation is needed before clinical use. Second, SICM status was determined at the episode level, whereas the model used the first examination within 24 h of sepsis onset; in some patients the defining abnormality may already have been present on that examination, and we did not separately evaluate prediction of later dysfunction. Third, there is no universally accepted diagnostic standard for SICM, and the development and external cohorts used different reference definitions. Fourth, the external cohort was small, which limits the precision of the performance and calibration estimates. Finally, only A4C cine loops were used, and we evaluated discrimination and physiological interpretability rather than clinical impact. Additional views, Doppler measurements, or strain may improve phenotypic characterization, and prospective studies are needed to determine whether model-assisted identification of SICM improves patient care [30].

538

539 **Conclusions**

540

541 We developed and externally evaluated a physiology-informed deep-learning framework
542 for early identification of SICM from apical four-chamber echocardiographic cine loops.

543 The model combined automated cardiac-cycle alignment, phase-aware spatiotemporal
544 representation learning, LV area-derived physiological information, and multi-cycle
545 aggregation. The model outperformed direct EchoJEPA-based classification, showed
546 consistent discrimination in patient-level fivefold cross-validation, and maintained strong
547 discrimination with good calibration in the independent MIMIC-IV-ECHO cohort.

548 Physiological and attention analyses showed that the model identified abnormal motion
549 patterns across different cardiac phases, spatial regions, and multiple cardiac cycles.

550 These abnormalities contributed to the identification of SICM. These findings support
551 further prospective evaluation of physiology-informed echocardiographic video analysis
552 for early identification of SICM.

553

554 **List of abbreviations**

555

556 **A4C:** Apical four-chamber

557 **APACHE II:** Acute Physiology and Chronic Health Evaluation II

558 **AUROC:** Area under the receiver operating characteristic curve

559 **CI:** Confidence interval

560 **cTnT:** Cardiac troponin T

561 **ED:** End diastole

562 **ES:** End systole

563 **hs-cTnT:** High-sensitivity cardiac troponin T

564 **ICU:** Intensive care unit

565 **LV:** Left ventricular

566 **LV-FAC:** Left ventricular fractional area change

567 **LVEF:** Left ventricular ejection fraction

568 **MIL:** Multiple-instance learning

569 **MIMIC-IV:** Medical Information Mart for Intensive Care IV

570 **MIMIC-IV-ECHO:** Medical Information Mart for Intensive Care IV Echocardiography

571 **ROC:** Receiver operating characteristic

572 **SICM:** Sepsis-induced cardiomyopathy

573 **SOFA:** Sequential Organ Failure Assessment

574 **TAPSE:** Tricuspid annular plane systolic excursion

575

576 **Declarations**

577 **Ethics approval and consent to participate**

578

579 This study was approved by the Ethics Committee of the First Affiliated Hospital with

580 Nanjing Medical University (approval No. 2026-SR-0215) and the Ethics Committee of

581 the Second Affiliated Hospital of Harbin Medical University (approval No. KY2026-036).

582 The requirement for informed consent was waived by both ethics committees because of

583 the retrospective nature of the study. This study was conducted in accordance with the

584 principles of the Declaration of Helsinki.

585

586 The external validation cohort was derived from MIMIC-IV v3.1 and MIMIC-IV-ECHO

587 v1.0.1, which contain de-identified patient data. Access to these databases was obtained

through PhysioNet by a credentialed investigator who completed the required training and agreed to the PhysioNet Credentialed Health Data Use Agreement. The collection and creation of the MIMIC-IV research resource were reviewed by the Institutional Review Board of Beth Israel Deaconess Medical Center, which granted a waiver of informed consent. The MIMIC-IV-ECHO project was approved by the Institutional Review Boards of Beth Israel Deaconess Medical Center and the Massachusetts Institute of Technology, and the requirement for individual patient consent was waived because all protected health information was de-identified and the project did not affect clinical care.

Consent for publication

Not applicable.

Availability of data and materials

The clinical and echocardiographic data from the two-center development cohort are not publicly available because they contain potentially sensitive patient information and are subject to institutional and ethical restrictions. De-identified data may be available from the corresponding author upon reasonable request and subject to approval by the relevant institutions and ethics committees.

The external validation data were obtained from MIMIC-IV v3.1 and MIMIC-IV-ECHO v1.0.1, which are available through PhysioNet to credentialed users who complete the required training and agree to the PhysioNet Credentialed Health Data Use Agreement.

613

614 The source code used for data preprocessing, model development, cross-validation, and
615 external evaluation is publicly available at <https://github.com/DrHan7/sicm-echo>.

616

617 **Competing interests**

618

619 The authors declare that they have no competing interests.

620

621 **Funding**

622

623 This work was supported by the National Natural Science Foundation of China (No.
624 82470496), the Regional United Key Program of the National Natural Science
625 Foundation of China (No. U25A2010), the National Key Research and Development
626 Program of China (No. 2023YFA1801900), the Outstanding Youth Project of the
627 Heilongjiang Natural Science Foundation (No. JQ2024H004), and the Research and
628 Practice Innovation Plan for Graduate Students in Jiangsu Province (SJCX24_0782).

629

630 **Authors' contributions**

631

632 YZ contributed to the study design, developed the code, performed the statistical and
633 computational analyses, and drafted the manuscript. QW critically revised the
634 manuscript for important intellectual content. JW contributed to data collection and
635 assisted with code revision. YH conceived and supervised the study and critically
636 revised the manuscript. All authors read and approved the final manuscript.

637

Acknowledgements

Not applicable.

References

1. Hasegawa D, Ishisaka Y, Maeda T, Prasitlumkum N, Nishida K, Dugar S, et al. Prevalence and prognosis of sepsis-induced cardiomyopathy: a systematic review and meta-analysis. *J Intensive Care Med.* 2023;38(9):797-808. doi:10.1177/08850666231180526.
2. Sato R, Sanfilippo F, Lanspa M, Duggal A, Dugar S. Sepsis-induced cardiomyopathy: mechanism, prevalence, assessment, prognosis, and management. *Chest.* 2025;168(6):1383-1394. doi:10.1016/j.chest.2025.08.013.
3. Lin YM, Lee MC, Toh HS, Chang WT, Chen SY, Kuo FH, et al. Association of sepsis-induced cardiomyopathy and mortality: a systematic review and meta-analysis. *Ann Intensive Care.* 2022;12(1):112. doi:10.1186/s13613-022-01089-3.
4. Tescione M, Piccolo A, Vadalà EG, Posteraro AG, Franzutti C, Macheda M, et al. How to do echo in septic cardiomyopathy: a consensus statement of the Italian Society of Echocardiography and Cardiovascular Imaging. *J Cardiovasc Echogr.* 2025;35(3):307-313. doi:10.4103/jcecho.jcecho_127_25.
5. Evans L, Rhodes A, Alhazzani W, Antonelli M, Coopersmith CM, French C, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. *Intensive Care Med.* 2021;47(11):1181-1247. doi:10.1007/s00134-021-06506-y.

6. Thygesen K, Alpert JS, Jaffe AS, Chaitman BR, Bax JJ, Morrow DA, et al. Fourth universal definition of myocardial infarction (2018). *J Am Coll Cardiol*. 2018;72(18):2231-2264. doi:10.1016/j.jacc.2018.08.1038.
7. Pruszczyk A, Zawadka M, Andruszkiewicz P, LaVia L, Herpain A, Sato R, et al. Mortality in patients with septic cardiomyopathy identified by longitudinal strain by speckle tracking echocardiography: an updated systematic review and meta-analysis with trial sequential analysis. *Anaesth Crit Care Pain Med*. 2024;43(2):101339. doi:10.1016/j.accpm.2023.101339.
8. Ouyang D, He B, Ghorbani A, Yuan N, Ebinger J, Langlotz CP, et al. Video-based AI for beat-to-beat assessment of cardiac function. *Nature*. 2020;580(7802):252-256. doi:10.1038/s41586-020-2145-8.
9. Sahashi Y, Ouyang D, Okura H, Kagiya N. AI-echocardiography: current status and future direction. *J Cardiol*. 2025;85(6):458-464. doi:10.1016/j.jjcc.2025.02.005.
10. Vrudhula A, Duffy G, Vukadinovic M, Liang D, Cheng S, Ouyang D. High-throughput deep learning detection of mitral regurgitation. *Circulation*. 2024;150(12):923-933. doi:10.1161/CIRCULATIONAHA.124.069047.
11. Christensen M, Vukadinovic M, Yuan N, Ouyang D. Vision-language foundation model for echocardiogram interpretation. *Nat Med*. 2024;30(5):1481-1488. doi:10.1038/s41591-024-02959-y.
12. Kim S, Jin P, Song S, Chen C, Li Y, Ren H, et al. EchoFM: foundation model for generalizable echocardiogram analysis. *IEEE Trans Med Imaging*. 2025;44(10):4049-4062. doi:10.1109/TMI.2025.3580713.
13. Vukadinovic M, Chiu IM, Tang X, Yuan N, Chen TY, Cheng P, et al. Comprehensive echocardiogram evaluation with view primed vision language AI. *Nature*. 2026;650(8103):970-977. doi:10.1038/s41586-025-09850-x.

14. Munim A, Fallahpour A, Szasz T, Attarpour A, Jiang R, Sooriyakanthan B, et al. EchoJEPA: a latent predictive foundation model for echocardiography. arXiv. 2026. doi:10.48550/arXiv.2602.02603.
15. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, et al. The third international consensus definitions for sepsis and septic shock (Sepsis-3). JAMA. 2016;315(8):801-810. doi:10.1001/jama.2016.0287.
16. Nagueh SF, Smiseth OA, Appleton CP, Byrd BF 3rd, Dokainish H, Edvardsen T, et al. Recommendations for the evaluation of left ventricular diastolic function by echocardiography: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. J Am Soc Echocardiogr. 2016;29(4):277-314. doi:10.1016/j.echo.2016.01.011.
17. Lang RM, Badano LP, Mor-Avi V, Afilalo J, Armstrong A, Ernande L, et al. Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. J Am Soc Echocardiogr. 2015;28(1):1-39.e14. doi:10.1016/j.echo.2014.10.003.
18. Gow B, Pollard T, Greenbaum N, Moody B, Han A, Waks JW, et al. MIMIC-IV-Echo: Echocardiogram Matched Subset (version 1.0.1). PhysioNet. 2026. <https://doi.org/10.13026/307c-mr50>.
19. Johnson AEW, Bulgarelli L, Shen L, Gayles A, Shammout A, Horng S, et al. MIMIC-IV, a freely accessible electronic health record dataset. Sci Data. 2023;10(1):1. doi:10.1038/s41597-022-01899-x.
20. Lau ES, Di Achille P, Kopparapu K, Andrews CT, Singh P, Reeder C, et al. Deep learning-enabled assessment of left heart structure and function predicts

711 cardiovascular outcomes. *J Am Coll Cardiol.* 2023;82(20):1936-1948.
 712 doi:10.1016/j.jacc.2023.09.800.

713 21. Aissaoui N, Boissier F, Chew M, Singer M, Vignon P. Sepsis-induced
 714 cardiomyopathy. *Eur Heart J.* 2025;46(34):3339-3353.
 715 doi:10.1093/eurheartj/ehaf340.

716 22. Dugar S, Sato R, Chawla S, You JY, Wang X, Grimm R, et al. Is left ventricular
 717 systolic dysfunction associated with increased mortality among patients with
 718 sepsis and septic shock? *Chest.* 2023;163(6):1437-1447.
 719 doi:10.1016/j.chest.2023.01.010.

720 23. Dugar S, Siuba MT, Sacha GL, Sato R, Moghekar A, Collier P, et al.
 721 Echocardiographic profiles and hemodynamic response after vasopressin
 722 initiation in septic shock: a cross-sectional study. *J Crit Care.* 2023;76:154298.
 723 doi:10.1016/j.jcrc.2023.154298.

724 24. Salih A, Boscolo Galazzo I, Gkontra P, Lee AM, Lekadir K, Raisi-Estabragh Z, et al.
 725 Explainable artificial intelligence and cardiac imaging: toward more interpretable
 726 models. *Circ Cardiovasc Imaging.* 2023;16(4):e014519.
 727 doi:10.1161/CIRCIMAGING.122.014519.

728 25. Collins GS, Dhiman P, Ma J, Schlusser MM, Archer L, Van Calster B, et al.
 729 Evaluation of clinical prediction models (part 1): from development to external
 730 validation. *BMJ.* 2024;384:e074819. doi:10.1136/bmj-2023-074819.

26. Riley RD, Archer L, Snell KIE, Ensor J, Dhiman P, Martin GP, et al. Evaluation of clinical prediction models (part 2): how to undertake an external validation study. *BMJ*. 2024;384:e074820. doi:10.1136/bmj-2023-074820.
27. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*. 2024;385:e078378. doi:10.1136/bmj-2023-078378.
28. Yang X, Sun W, Chen K, Wang X. Establishment and validation of a critical care echocardiography-based predictive model for sepsis-induced cardiomyopathy: a prospective cohort study. *J Crit Care*. 2025;88:155066. doi:10.1016/j.jcrc.2025.155066.
29. Pinsky MR, Bedoya A, Bihorac A, Celi L, Churpek M, Economou-Zavlanos NJ, et al. Use of artificial intelligence in critical care: opportunities and obstacles. *Crit Care*. 2024;28(1):113. doi:10.1186/s13054-024-04860-z.
30. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, et al. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. *Nat Med*. 2022;28(5):924-933. doi:10.1038/s41591-022-01772-9.

Figure legends

Fig. 1 Study design and cohort selection. The study comprised a retrospective two-center development cohort and an independent external validation cohort derived from

MIMIC-IV-ECHO linked with MIMIC-IV. The model was developed in the two-center cohort and locked before external validation. A4C, apical four-chamber; cTnT, cardiac troponin T; SICM, sepsis-induced cardiomyopathy.

Fig. 2 Cardiac-cycle extraction and physiological alignment of echocardiographic cine loops. LV segmentation generated an area-time trajectory from each A4C cine loop. ED and ES were identified from the trajectory, and adjacent ED-to-ED intervals were extracted as individual cardiac cycles. Cycles of different durations were temporally resampled to a standardized 16-frame sequence spanning ED, systole, ES, diastole, and the subsequent ED. The highlighted interval indicates the representative ED-to-ED cycle shown below. A4C, apical four-chamber; ED, end diastole; ES, end systole; LV, left ventricular.

Fig. 3 Architecture and training strategy of the physiology-informed deep-learning framework. Standardized A4C cardiac cycles were encoded by EchoJEPA to generate spatiotemporal representations, followed by phase-aware spatial and temporal attention. In parallel, the LV area-time trajectory was encoded as a physiological representation of cardiac contraction and filling. Visual and physiological representations were fused at the cardiac-cycle level, and multiple cycles from the same examination were aggregated by attention-based MIL to generate the patient-level SICM probability. Self-supervised adaptation combined masked latent prediction within individual cycles with cross-cycle consistency across heartbeats from the same examination. Supervised training was subsequently performed in a frozen-backbone stage followed by partial fine-tuning of the EchoJEPA encoder. A4C, apical four-chamber; LV, left ventricular; MIL, multiple-instance learning; SICM, sepsis-induced cardiomyopathy.

778

779 **Fig. 4 Model performance and external validation.** a Receiver operating characteristic
780 (ROC) curves of the physiology-informed model in the development and external
781 validation cohorts. b Calibration of predicted SICM probabilities in the external validation
782 cohort; the shaded region represents pointwise 95% confidence intervals. c Decision
783 curve analysis in the external validation cohort comparing the physiology-informed
784 model with the EchoJEPA baseline. d ROC curves comparing the physiology-informed
785 model with the EchoJEPA baseline in the external validation cohort; AUROCs were
786 compared using the paired DeLong test. e Confusion matrix of the physiology-informed
787 model in the external validation cohort. f AUROC estimates with 95% confidence
788 intervals for the physiology-informed model and EchoJEPA baseline in the development
789 and external validation cohorts. AUROC, area under the receiver operating characteristic
790 curve; CI, confidence interval; ROC, receiver operating characteristic.

791

792 **Fig. 5 Physiological characteristics of aligned cardiac cycles.** a–c Patient-level
793 distributions of LV fractional area change, normalized systolic contraction rate, and
794 normalized diastolic filling rate in patients with SICM and sepsis-only controls. Each
795 point represents one patient after averaging values across available cardiac cycles. d
796 Mean normalized LV area trajectory across the standardized 16-frame cardiac cycle;
797 shaded regions indicate 95% confidence intervals. ED, end diastole; LV, left ventricular.

798

799 **Fig. 6 Spatial, temporal, and multi-cycle attention patterns of the physiology-**
800 **informed model.** a, b Representative sepsis-only and SICM examinations, respectively.
801 Raw A4C images and corresponding spatial-attention maps are shown at ED, systole,
802 ES, and late diastole. Temporal-attention weights across the eight temporal tokens are

shown below each examination. Spatial-attention maps in a and b use the same normalization scale. c Multi-cycle analysis of the representative SICM examination shown in b. The three cycles with the highest MIL attention weights are shown in their original order, together with MIL attention weights, LV fractional area change, and cycle length. A4C, apical four-chamber; ED, end diastole; ES, end systole; LV-FAC, left ventricular fractional area change; MIL, multiple-instance learning; SICM, sepsis-induced cardiomyopathy.

Additional files

Additional file 1 (.docx): Supplementary tables and figures. Contains Supplementary Tables S1–S3 and Supplementary Figs. S1–S2.

Supplementary Files

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

- [Additionalfile1.docx](#)