A GCE Framework for Interpretable Nonlinear Hazard Modeling in Cardiac Sarcoma Survival

Muhammad Shoaib Kareem , Madiha Amjad , Saba Aslam , Abdur Rasool * , Mutiullah Jamil , Hazrat Ali *

Posted Date: 3 June 2026

doi: 10.20944/preprints202606.0294.v1

Keywords: cardiac sarcoma; survival prediction; GRU; Cox Proportional Hazards; interpretable AI; SEER database

 doi:10.20944/preprints202606.0294.v1

Article

A GCE Framework for Interpretable Nonlinear Hazard Modeling in Cardiac Sarcoma Survival

Muhammad Shoaib Kareem [1], Madiha Amjad [1], Saba Aslam [2], Abdur Rasool [3,4,*], Mutiullah Jamil [1] and Hazrat Ali [5,*]

[1] Department of Information Technology, Khawaja Fareed University of Engineering and Technology, Rahim Yar Khan, 64200, Pakistan
[2] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, 518055, China
[3] Department of Information and Computer Sciences, University of Hawaii at Manoa, Honolulu, HI 96822, USA
[4] Department of Statistics and Data Science, Southern University of Science and Technology, Shenzhen, 518055, China
[5] Division of Computing Science and Mathematics, University of Stirling, Stirling, FK9 4LA, United Kingdom
* Correspondence: abdur@hawaii.edu (A.R.); hazrat.ali@live.com (H.A.)

Simple Summary

Cardiac sarcoma is a rare and highly aggressive cancer associated with poor survival outcomes and limited treatment guidance. Accurate prediction of patient survival is essential for improving clinical decision-making; however, conventional statistical approaches often fail to capture the complex, nonlinear patterns of cancer progression, and many artificial intelligence methods lack clinical interpretability. In this study, we propose a novel interpretable artificial intelligence framework that combines deep learning with statistical survival modeling to improve prognostic prediction for patients with cardiac sarcoma. Using data from SEER program, the proposed framework achieved superior predictive performance compared with existing survival models and identified clinically meaningful risk factors for patient outcomes. By integrating prediction accuracy with transparent interpretation, this framework may support earlier risk stratification, personalized treatment planning, and future development of clinically reliable artificial intelligence systems for rare cancers.

Abstract

Cardiac sarcoma is a rare and aggressive malignancy with a poor prognosis, where accurate survival prediction is critical yet hindered by conventional models' inability to capture nonlinear temporal dynamics and lack of interpretability. This study presents the GRU-CoxPH Ensemble (GCE), a weighted late-fusion interpretable framework for nonlinear hazard modeling that integrates Gated Recurrent Unit (GRU) networks with the Cox Proportional Hazards (CoxPH) model to address these limitations. From SEER data, 27 features were selected from 41 variables using the Least Absolute Shrinkage and Selection Operator (LASSO) and Random Survival Forests (RSF). The GCE framework combines a GRU network (capturing nonlinear temporal hazard patterns) with a CoxPH baseline (providing statistical anchoring) via weighted late-fusion. Strict exclusion of outcome-related variables prevented target leakage. The GCE framework achieved a mean C-index of 0.9830 and IBS of 0.03958 across a 10-fold cross-validation, outperforming standalone GRU (C-index 0.9345, IBS 0.05105) and classical CoxPH (C-index 0.8842, IBS 0.03280). Shapley Additive exPlanations (SHAP) analysis provided interpretable insights into feature importance, confirming clinical relevance. The GCE framework delivers robust, interpretable nonlinear hazard modeling for cardiac sarcoma survival prediction by capturing temporal dynamics while maintaining statistical transparency—addressing key limitations of conventional methods in small, rare-disease cohorts.

Keywords: cardiac sarcoma; survival prediction; GRU; Cox Proportional Hazards; interpretable AI; SEER database

 doi:10.20944/preprints202606.0294.v1

1. Introduction

Cardiac sarcoma represents a rare and highly aggressive malignancy and poses significantly greater threats to patient survival compared to other cardiac disorders [1]. This neoplasm is characterized by rapid progression and a dismal prognosis, with median survival rates often ranging from seven months to three years post-diagnosis [2]. The inherent challenges in managing cardiac sarcoma stem from its elusive early detection, heterogeneous histopathological presentations, and the limited efficacy of available therapies, which are frequently confounded by the tumor's invasive nature and propensity for metastasis. Figure 1 illustrates the localization of cardiac tumors, with (a) depicting the blob-like structure of angiosarcoma and (b) its corresponding texture [3].

The scarcity of comprehensive patient records exacerbates the difficulty in evaluating therapeutic impacts on survival outcomes [4]. Accurate survival prediction is pivotal in managing cardiac sarcoma and clinical research. Electronic health data from the SEER program serves as a key resource, compiling cancer incidence and survival statistics from U.S. registries [5]. Selecting predictive models from SEER data remains challenging, especially for integrating features to forecast outcomes in rare malignancies such as cardiac sarcoma [4,8]. The advent of deep learning (DL) and machine learning (ML) models has transformed prognostic analytics by capturing subtle patterns in textual, numerical, and visual data overlooked by traditional methods [11]. DL/ML applications on SEER data show superior performance across various cancers and indicate potential, such as in cardiac tumors [12]. CNN-RNN architectures, for example, enable multimodal integration, enhancing predictive accuracy and forecasts by combining imaging and temporal features. Such models analyze interactive effects in large datasets, providing finer-grained outcome analyses than conventional statistics [15,20].

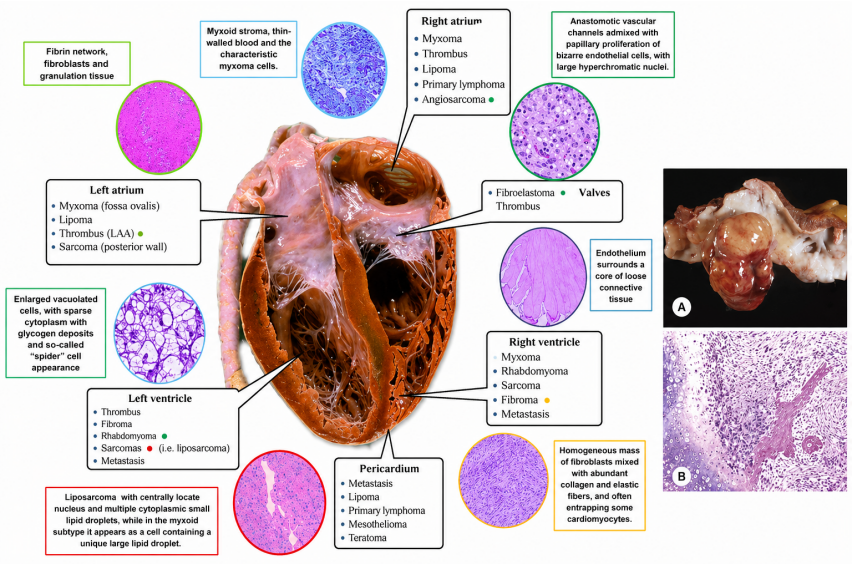


Figure 1. Representative images of cardiac angiosarcoma. Left: macroscopic appearance showing a nodular, hemorrhagic lesion in the heart. Right: microscopic features showing irregular vascular spaces lined by atypical endothelial cells [3]

This is particularly relevant in cardio-oncology, where ML assesses cancer progression alongside cardiovascular comorbidities, the leading mortality cause in survivors of breast, prostate, or bladder cancers [20,24]. Despite these advancements, significant gaps remain in cardiac sarcoma prognostication. The disease's rarity yields small cohorts (e.g., fewer than 1,000 cases in national databases), causing data sparsity that exacerbates overfitting, bias, and poor generalizability [2]. Traditional approaches, such as the CoxPH model, dominate but fail to capture complex nonlinear relationships in heterogeneous survival data [31]. SEER-based studies on related cancers often report modest C-index values (0.77–0.87) and overlook comprehensive metrics such as IBS, thereby limiting clinical utility [21,31]. The aggressive biology of cardiac sarcoma, combined with treatment-related cardiotoxi-

 doi:10.20944/preprints202606.0294.v1

city and competing risks (e.g., cardiovascular mortality)—requires models handling censored data, temporal dependencies, and interpretability [8].

Three key problems remain unaddressed: (1) conventional CoxPH cannot model nonlinear hazard dynamics [38,63]; (2) DL models lack interpretability and overfit in small cohorts [28,56]; (3) no systematic framework exists that jointly provides nonlinear hazard modeling, statistical transparency, and clinical explainability for rare-cancer survival prediction [43,60].

To address these challenges, this study presents the GRU-CoxPH Ensemble (GCE), a weighted late-fusion interpretable framework for nonlinear hazard modeling in cardiac sarcoma survival. The GCE integrates a GRU network (capturing complex temporal hazard patterns) with a CoxPH baseline (providing statistical anchoring) via optimized convex weighting.

The specific objective of this study is to quantify whether a weighted late-fusion ensemble of GRU and CoxPH (GCE) improves discrimination (C-index) and calibration (IBS) over standalone DL and traditional survival models for cardiac sarcoma prognosis using SEER data, while providing clinically interpretable risk factors via SHAP analysis.

The following are the major contributions of this work:

- Novel ensemble architecture: A weighted late-fusion of GRU and CoxPH that uniquely combines deep temporal learning with statistical survival analysis—unlike prior work that uses either deep OR statistical models, not both synergistically.
- Interpretability in nonlinear hazard modeling: The GCE captures time-varying risk patterns while providing SHAP-based feature importance and CoxPH-anchored transparency—addressing the black-box criticism of DL in clinical settings.
- Robust rare-disease validation: Rigorous feature engineering (LASSO, RSF, PCA) and 10-fold CV with strict leakage prevention on a small SEER cohort (n=727), achieving strong performance (C-index 0.9830, IBS 0.03958) with clinical actionability.

The proposed framework advances medical informatics with accurate, interpretable AI for personalized survival prediction in rare cardiac sarcoma using SEER data. This paper is structured as follows: Section 1 provides the introduction. Section 2 presents the literature review. Section 3 details the methodology, including data, variables, and model development. Section 4 presents the results. Section 5 reports the discussion and limitations. Section 6 provides the conclusion.

2. Literature Review

2.1. Cancer Survival Prediction Using AI

The Surveillance, Epidemiology, and End Results (SEER) database has been widely used for survival prediction in cancers like sarcomas, cardiac tumors, and gastrointestinal mesenchymal tumors, covering approximately 34% of the U.S. population with demographic, clinical, and follow-up data updated yearly [22,33]. Linked datasets such as SEER-Medicare provide additional treatment details [27,34]. Other datasets include the National Cancer Database (NCDB) for cardiac angiosarcoma [8] and CLARO for lung cancer [32]. In cardiac malignancies, SEER supports studies on primary sarcoma, lymphoma, and angiosarcoma [2,8]. For broader cancers, SEER aids predictions in hepatocellular carcinoma [27,28], gastric adenocarcinoma [13], and glioblastoma [36]. DeepSurv extends CoxPH for SEER data in gastric adenocarcinoma and melanoma, achieving C-indexes of 0.825–0.871 [13,37]. MTLR variants perform in breast and hepatocellular carcinoma with C-indexes around 0.771–0.824 [28,38] Studies enhance prognostication using metrics such as C-index, Integrated Brier Score (IBS), and AUC to inform clinical decision-making.

2.2. Challenges in SEER-Based Prognostic Modeling

SEER studies face limitations in data quality, including retrospective biases and missing details such as comorbidities, treatment dosages, genetic information, or lifestyle factors (e.g., smoking/BMI captured via inaccurate ICD codes) [7,8,22,34,39]. SEER-Medicare focuses on patients over 65 years old, reducing generalizability to younger or uninsured populations, while geographic coverage and

 doi:10.20944/preprints202606.0294.v1

rare event undercounting (e.g., sudden cardiac deaths) introduce further inaccuracies [25,41,65]. These gaps hinder robust evaluations in cardiac lymphoma and angiosarcoma [2,33]. Traditional statistical models such as CoxPH, logistic regression, and Kaplan-Meier fail to capture nonlinear interactions in survival data [25,37]. DL methods like DeepSurv and hybrid models (CNN-LSTM, Deep-CRMTLR) improve multimodal predictions in lung, breast, and hepatocellular carcinoma but often yield modest C-indexes (0.77–0.87) and frequently omit IBS calibration [15–17]. In cardio-oncology, models may overlook temporal dependencies and interpretability in competing risks for rare tumors [7,22,23]. Validation typically relies on basic metrics (AUC, $p < 0.05$) without rigorous cross-validation or ensemble evaluation [7,26]. Furthermore, we have summarized and compared the studies in Table 1. Our proposed GCE framework, combined with LASSO, RSF, and Principal Component Analysis (PCA) for feature selection, 10-fold cross-validation, high C-index/IBS performance (>0.98/0.03), and SHAP-based interpretability, addresses these challenges for cardiac sarcoma.

doi:10.20944/preprints202606.0294.v1

Table 1. Summary of selected studies on survival prediction in various cancers using SEER and related datasets, highlighting datasets, methods, and key outcomes.

Author (Year)	Disease / Dataset	Features	Model	Evaluation Parameter / Findings	Challenges
[43] (2023)	Gastric adenocarcinoma / SEER	Clinical and demographic variables	DeepSurv, CoxPH, RSF	C-index 0.825–0.871, IBS reported; Improved nonlinear handling over CoxPH	Modest accuracy (below 0.90); Limited interpretability of DL predictions; No ensemble fusion; Overlooks temporal dependencies in rare cohorts
[28] (2024)	Hepatocellular carcinoma / SEER	Tumor characteristics, demographics, treatments	N-MTLR, CoxPH, DeepSurv, RSF	C-index 0.824, IBS; Better than baselines in high-dimensional data	Accuracy below 90%; No hybrid statistical-DL integration; Ignores competing risks like cardiovascular mortality; Lacks SHAP-based feature insights
[38] (2025)	Breast cancer / SEER	Clinical, imaging, demographics	N-MTLR, CoxPH, RSF	C-index 0.771–0.821, IBS; CoxPH strong but DL adds nonlinearity	Low overall accuracy; No temporal modeling (GRU/LSTM); Limited to non-rare tumors; Absence of feature engineering (PCA/LASSO) for small datasets
[8] (2023)	Cardiac angiosarcoma / NCDB	Histology, stage, treatments	Cox regression	95% CI statistical analysis; Identified geographic variations	Descriptive only; No predictive modeling; Low focus on rare cardiac sarcoma; No DL for nonlinear patterns; Generalizability issues
[35] (2021)	Primary cardiac sarcoma / SEER	Demographics, tumor size, survival	Univariate/ multivariate regression	Regression analysis; Prognostic factors identified	Analysis-focused, not prediction; No DL/ML for complexity; Missing interpretability; Fails to address SEER's data sparsity
[36] (2024)	Glioblastoma / SEER	Clinical features	XGBoost, AdaBoost, DT, KNN, RF, DNN	MSE, RMSE (%) 90.25; ML vs DL comparison	Uses non-survival metrics; No hybrids for temporal data; Overfitting in small cohorts; Lacks transparency for clinical use
[2] (2020)	Primary cardiac lymphoma / SEER	Age, histology, survival	Kaplan-Meier, statistical analysis	IQR, survival curves; Descriptive trends	Descriptive only; No predictive models; Ignores nonlinear interactions; Limited to lymphoma, not sarcoma; No validation for robustness
[32] (2024)	Lung cancer / CLARO	Clinical features	AI model	C-index = 80.72	Low accuracy; No ensemble or hybrid; Limited interpretability; Dataset-specific, not generalizable
[5] (2023)	Actigraphy Data & Clinical Information	Clinical Information	DL models: LSTM, BiLSTM, GRU, RNN	KPS, Palliative Performance Index (PPI) = 0.89	Few researchers used this; No SEER integration; Lacks fusion with statistical models; No SHAP for feature importance

 doi:10.20944/preprints202606.0294.v1

3. Materials and Methods

3.1. Data Collection and Patient Selection

This retrospective cohort study was conducted using data from the Surveillance, Epidemiology, and End Results (SEER) program available online: (https://seer.cancer.gov/) [22]. Given the extreme rarity of primary cardiac sarcoma, SEER represents one of the few sources with sufficient cases for meaningful statistical analysis [41]. Data extraction was performed with SEER*Stat software (version 8.4.1; November 2024 submission, covering 21 registries excluding Illinois, diagnosis years 2000–2022). From an initial pool of 16,057,864 records, 727 cases met eligibility criteria after rigorous filtering

To prevent target leakage, all outcome-related variables (Survival months, Vital status recode, Year of follow-up recode, SEER cause-specific death classification, and COD to site recode) were strictly excluded from the feature set before any preprocessing, feature selection, or model training. Feature engineering (LASSO, RSF, and PCA) and model development were performed exclusively on baseline covariates available at the time of diagnosis. Ten-fold stratified cross-validation was conducted with preprocessing pipelines fitted only on the training folds to further ensure no information from the test set influenced model development. Owing to the public, de-identified nature of SEER data, the study was exempt from institutional review board approval [11,18].

3.2. Feature Selection

Candidate variables (n = 41) were initially retrieved from SEER, spanning four domains: demographics (age at diagnosis recoded $\leq$65 vs. >65 years, sex, race/ethnicity, county-level median household income quartiles), tumor characteristics (ICD-O-3 histologic type and behavior, grade, size recoded $\leq$4 cm vs. >4 cm, SEER historic stage, primary site confirmation), treatment (surgery type/code, year of diagnosis), and survival outcomes (vital status recode, survival months, SEER cause-specific death classification). Variables with excessive missingness (>20%, e.g., detailed chemotherapy/radiation fields, comorbidities) were excluded *a priori* to avoid imputation bias in a small cohort. All variables were strictly defined at the time of diagnosis (time-zero) to ensure consistency. Outcome-related variables were removed before preprocessing and feature selection so that only baseline clinical information contributes to model development. Feature selection was performed using LASSO regression and RSF [12].

$$\min_{\beta}\left[-\frac{1}{n}\sum_{i=1}^{n}\delta_i\left(x_i^T\beta-\log\sum_{j\in R(t_i)}\exp\left(x_j^T\beta\right)\right)+\lambda\|\beta\|_1\right] \quad (1)$$

where λ was tuned via 10-fold cross-validation to balance sparsity and predictive performance [44,45].

RSF provided variable importance rankings based on permutation importance and out-of-bag error, enabling robustness against nonlinear relationships and correlated predictors [47]. To further strengthen the reliability of selected features, a simple stability measure was introduced to quantify how consistently each feature is selected across folds:

$$SF_j=\frac{1}{K}\sum_{k=1}^{K}I_j^{(k)} \quad (2)$$

where SF_j represents the selection frequency of feature j, K is the number of folds, and $I_j^{(k)}$ equals 1 if the feature is selected in fold k, otherwise 0.

Features with higher SF_j values were considered more stable and reliable for downstream modeling. All retained variables were organized into clearly defined categories (demographic, tumor-specific, and treatment-related), ensuring that each feature contributes uniquely to the analysis and maintains clear clinical interpretation. Consensus features ($n = 27$) were retained for downstream modeling, prioritizing those consistently ranked high across both methods.

Feature selection was integrated within the cross-validation pipeline, where it was applied only on training folds. This preserves the independence of validation data and ensures reliable performance estimation. This structured approach produces a stable and compact feature set, improving model robustness while maintaining interpretability in the context of rare-disease survival prediction.

 doi:10.20944/preprints202606.0294.v1

3.3. Proposed GCE Framework

The extreme rarity of primary cardiac sarcoma and limited sample size ($n = 727$ after stringent filtering) pose fundamental challenges to survival analysis. Traditional models such as CoxPH assume linear covariate effects, which are frequently violated in risk profiles characteristic of rare malignancies. DL architectures are effective in capturing nonlinear relationships but may introduce instability and reduced interpretability when applied to small datasets.

To address these challenges, we propose a weighted late-interpretable GCE framework that integrates the strengths of GRU with the statistical structure of the CoxPH model. This design enables simultaneous modeling of nonlinear feature interactions while preserving clinically meaningful risk structure. The overall workflow is illustrated in Figure 2. The central component of the framework is the ensemble fusion strategy. After training individual models, a patient-specific risk score E_i is computed as:

$$E_i = w_{\text{CoxPH}} \cdot S_{\text{CoxPH},i} + w_{\text{GRU}} \cdot S_{\text{GRU},i} \tag{3}$$

where, $S_{\text{CoxPH},i}$ corresponds to the log-risk estimated from the Cox model, reflecting proportional hazard assumptions, while $S_{\text{GRU},i}$ represents the nonlinear risk embedding learned through sequential modeling. The weights w_{CoxPH} and w_{GRU} determine the relative contribution of each component and satisfy the constraint $w_{\text{CoxPH}} + w_{\text{GRU}} = 1$.

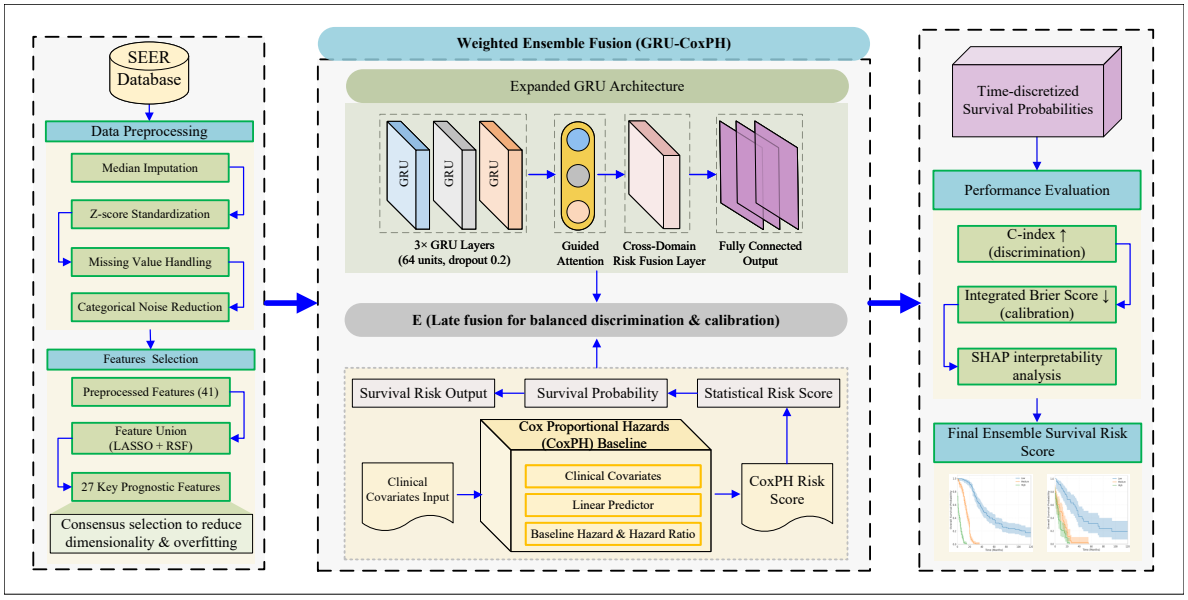


Figure 2. The schematic framework of the expanded GRU network within the GCE framework, SEER data extraction and cohort curation, feature selection, parallel training of survival learners under 10-fold CV, weighted late fusion of the most performant components, optimization of discrimination (C-index) and calibration (IBS), and post-hoc SHAP analysis for clinical-grade explainability. End-to-end design ensures robustness against data sparsity while maximizing predictive power and transparency.

To ensure that both models contribute comparably, the raw risk outputs are first transformed and normalized through the following formulation:

$$\tilde{S}_{m,i} = \frac{S_{m,i} - \mu_m}{\sigma_m}, \quad E_i = \sum_{m \in \{\text{CoxPH,GRU}\}} w_m \cdot \tilde{S}_{m,i} \tag{4}$$

where μ_m and σ_m denote the mean and standard deviation of model m computed from training data. This formulation ensures scale alignment and prevents dominance of one model due to magnitude differences. With empirically determined weights $w_{\text{CoxPH}} = 0.2$ and $w_{\text{GRU}} = 0.8$, obtained via grid-based optimization on validation folds.

 doi:10.20944/preprints202606.0294.v1

The ensemble mechanism operates by combining complementary learning representations. The CoxPH model provides a stable and interpretable estimation of baseline hazard trends, while the GRU captures complex nonlinear dependencies and implicit temporal patterns across patient features. The final score E_i therefore integrates both global statistical structure and local nonlinear variations in risk. From an operational perspective, the computation of E_i proceeds in three stages: (1) each model independently processes the input features and generates a risk score, (2) the scores are normalized to ensure consistency, and (3) the weighted combination produces a unified risk estimate. This stepwise design maintains separation between model-specific learning and final aggregation.

This formulation ensures stable predictions in small datasets by reducing sensitivity to individual model fluctuations. If one component produces high-variance outputs, the complementary model stabilizes the final prediction through weighted averaging. This late-fusion strategy differs from early fusion approaches that combine raw features or intermediate representations. Instead, it preserves model independence and allows each component to specialize in different aspects of the survival process.

The temporal backbone of the framework is an extended GRU architecture tailored for survival modeling.

Multi-scale temporal abstraction: Three stacked GRU layers (64 hidden units each, dropout = 0.2) enable hierarchical learning of short-term and long-term temporal dependencies.

Capsule-guided attention mechanism: Let $H = [h_1, \ldots, h_T]$ denote the sequence of hidden states. Capsule outputs are defined as:

$$u_j = \sum_{t=1}^{T} c_{jt}\, h_t, \quad c_{jt} = \frac{\exp\left(q_j^T h_t / \sqrt{d}\right)}{\sum_s \exp\left(q_j^T h_s / \sqrt{d}\right)} \tag{5}$$

This mechanism assigns adaptive importance to temporal states, enabling the model to focus on clinically relevant patterns while suppressing less informative signals.

The cross-domain risk fusion layer integrates deep and statistical components:

$$S(\tau \mid x) = \exp\left(-\int_0^\tau \lambda_0(u) \exp\left(f_{\mathrm{GRU}}(x,u) + \beta^\top x\right) du\right) \tag{6}$$

which is approximated as:

$$\hat{S}_{t,\tau} = \sigma\left(\rho(z(t)) \cdot S_{\mathrm{GRU}}(x,\tau) + (1 - \rho(z(t))) \cdot S_{\mathrm{CoxPH}}(\tau \mid x)\right) \tag{7}$$

where the gated function $\rho(z(t))$ dynamically adjusts the contribution of each component over time, enabling adaptive fusion based on temporal context and ensuring smooth survival probability estimation.

The complete framework integrates preprocessing, feature selection, model training, and ensemble fusion into a unified pipeline, ensuring consistent data handling and stable predictive behavior across validation folds. The complete workflow, including preprocessing steps such as z-score standardization and median imputation, is illustrated in Figure 2, while the detailed procedure is described in Algorithm 1.

3.4. Baseline and Core Models

Seven widely used survival models were trained in parallel for cardiac sarcoma survival prediction using the 27 selected features. CoxPH serves as the statistical baseline [39,49]. RSF is capable of nonlinear and complex features [50,51]. DeepSurv learns nonlinear hazard [53]. Multi-Task Logistic Regression (MTLR) analysis across discretized time [64]. Convolutional Neural Network (CNN) spatial feature to temporal modeling [16,54]. Long Short-Term Memory (LSTM) captures extended temporal patterns [56]. GRU constitutes the central component [58]. All models were implemented using lifelines, pycox, and PyTorch.

 doi:10.20944/preprints202606.0294.v1

Algorithm 1 GCE Framework for Survival Risk Prediction

Input: Dataset X, time T, event δ
Preprocess data (imputation, normalization, feature filtering)
Select features via LASSO + RSF $\rightarrow X^*$
Train CoxPH model $\rightarrow S_{\text{CoxPH},i}$
for each sample X_i^* **do**
    Generate sequence and train GRU $\rightarrow H_i$
    Apply attention (Eq. (3) or use last hidden state
    Compute $S_{\text{GRU},i}$
end for
for each patient i **do**
    Estimate survival probability using Eq. (5)
    Compute continuous risk score from survival output
end for
Compute ensemble risk score:
    $E_i = w_{\text{CoxPH}}\, S_{\text{CoxPH},i} + w_{\text{GRU}}\, S_{\text{GRU},i}$
Optimize weights via cross-validation (C-index, IBS)
Output: Final predictions E_i

3.5. Experimental Setup and Evaluation

This framework ensures reliable generalization and minimizes overfitting in this small and heavily censored rare-disease cohort. Models were evaluated using a repeated stratified k-fold cross-validation scheme. Detailed configurations, including model architectures, are provided in Table **??**. Mean performance and fold-to-fold variability were reported to assess model stability. For models requiring sequential input, 27 selected features were reorganized into pseudo-temporal sequences to enable the networks to learn potential dependencies among variables. Experiments were conducted on a GPU-accelerated system with PyTorch for deep components and lifelines, and Pycox for statistical baselines.

Table 2. Hyperparameters, training configurations, and evaluation settings for the GCE framework.

Configuration Group	Component / Parameter	Value / Setting
Data & Input	Dataset	SEER survival dataset
	Survival Time Variable	Survival months
	Event Indicator	Vital status recode (0 = censored, 5 = event)
	Value Handling	Median imputation (numeric)
	Random Seed	42
Validation Strategy	Cross-Validation	10-Fold Cross-Validation (shuffle = True)
	Train–Validation Split	80% Train, 20% Validation (within training fold)
	Stratification	Based on event status
Cox Proportional Hazards	Model Type	Baseline/Lifelines CoxPH
	Penalization	L2 penalizer = 0.1
	Output	Partial hazard scores (Risk)
Survival Model	Architecture–GRU	3-layer GRU
	Hidden Units	64 Units
	Dropout	0.2 (between layers)
	Temporal Modeling	Learns nonlinear feature interactions & latent survival dynamics
	Input Format	Feature vector reshaped to sequence
	Optimizer	Adam
	Epochs	50
	Batch Size	64 (Model training)
	Output	Fully Connected linear layer
Ensemble Strategy	Fusion Method	Weighted averaging
	Ensemble Weights	$0.2 \times$ CoxPH + $0.8 \times$ GRU
Evaluation Setup	Evaluation Time Grid	100 time points between min–max survival
	Baseline Survival	Kaplan–Meier estimator
Performance Metrics	Discrimination Metric	Concordance Index (C-index)
	Calibration Metric	Integrated Brier Score (IBS)
	Reporting	Per-fold and mean performance across 10-fold

 doi:10.20944/preprints202606.0294.v1

4. Results

4.1. Performance of GCE Framework

The mean performance across all folds is reported in Table 3. The proposed GCE framework (weighted integration of CoxPH at 0.2, and GRU at 0.8) delivered the highest mean C-index of 0.9830 with an IBS of 0.03958—higher than the standalone GRU-only mean C-index of 0.9345, and IBS of 0.05105, demonstrating outstanding discriminative capability but reflecting an excellent balance between discrimination and calibration. The LSTM model also showed strong results (mean C-index 0.9289, IBS 0.04985), whereas the CoxPH baseline trailed substantially (mean C-index 0.8842, IBS 0.03280).

This framework highlights the critical advantage of combining deep temporal learning (GRU) with statistical grounding (CoxPH) via weighted fusion—thereby addressing the prognostic limitations of conventional approaches in small, high-censoring, rare-tumor datasets. These findings position the GCE framework as the leading solution for accurate, reliable, and clinically interpretable survival prediction in primary cardiac sarcoma, significantly outperforming baseline and individual models while overcoming data sparsity and nonlinearity challenges.

Table 3. Performance of deep learning models and mean C-index and IBS across 10-fold cross-validation.

Evaluation Category	Model	C-index	IBS
Performance of Deep learning models	CNN [54]	0.884399	0.00585
	LSTM [56]	0.893473	0.00604
	GRU [58]	0.901388	0.00516
	Proposed Framework	0.936179	0.00417
Mean Score across 10-fold CV	CoxPH [49]	0.8842	0.03280
	LSTM [56]	0.9289	0.04985
	GRU [58]	0.9345	0.05105
	Proposed Framework	0.9830	0.03958

4.2. Impact of Feature Selection and Reduction Strategies

To address multicollinearity and improve model robustness in the limited SEER cohort, employing the proposed GCE framework. PCA was first applied to reduce the 27 features, detailed ratios in Table 4. Detailed reports of performance in Table 5 after using PCA-transformed features (20 components retained). CoxPH and DeepSurv showed modest improvements in C-index (0.8719 and 0.8649, respectively) and better calibration (IBS 0.03541 and 0.01463). RSF and MTLR, however, exhibited declines (C-index 0.7858 and 0.8381), suggesting PCA's linear transformation was less advantageous for inherently nonlinear methods.

CoxPH and RSF (C-index 0.31–0.75), indicating insufficient retention of predictive information. RSF-selected features improved RSF and DeepSurv (C-index up to 0.8255). The union set with DeepSurv achieved the highest C-index (0.856182) and lowest IBS (0.047134), detailed in Table 6. Comparative performance across feature selection strategies is visually shown in Figure 3, comparing the C-index and IBS across LASSO, RSF, and union sets for the four models (subfigures a and b).

The bar plots highlight the clear superiority of the union strategy, particularly for DeepSurv, where C-index increases substantially while IBS decreases markedly compared to LASSO or RSF alone. These findings demonstrate that combining LASSO's sparsity with RSF's importance ranking yields a more discriminative and well-calibrated feature set—directly improving input quality for the proposed GCE framework and addressing data sparsity/nonlinearity challenges in rare-disease survival modeling.

 doi:10.20944/preprints202606.0294.v1

Table 4. Explained variance ratios for the top 16 principal components from PCA on the 27 selected features.

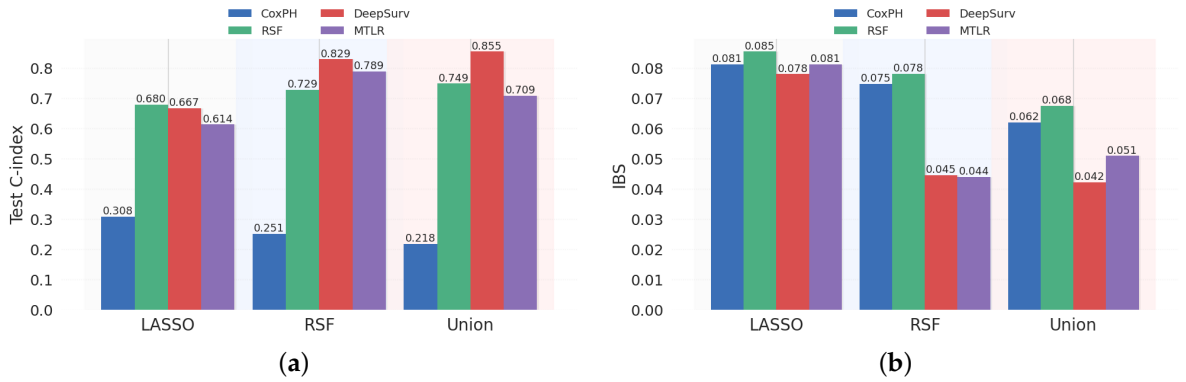
Principal Component	Explained Variance Ratio	Principal Component	Explained Variance Ratio
PC1	0.155156	PC9	0.039989
PC2	0.126531	PC10	0.039268
PC3	0.093936	PC11	0.036696
PC4	0.077580	PC12	0.036647
PC5	0.068787	PC13	0.034826
PC6	0.051185	PC14	0.030896
PC7	0.048991	PC15	0.027285
PC8	0.042818	PC16	0.079200

Table 5. Performance of classical and baseline models using PCA-transformed features (20 components) and 27 selected features.

Evaluation Category	Model	Test C-index	IBS
PCA-transformed features (20)	CoxPH [49]	0.8719	0.03541
	RSF [51]	0.7858	0.05762
	DeepSurv [53]	0.8649	0.01463
	MTLR [15]	0.8381	0.02344
	Proposed Framework	0.9830	0.03958
Selected features (27)	CoxPH [49]	0.8842	0.03280
	RSF [51]	0.8322	0.04830
	DeepSurv [53]	0.9061	0.01990
	MTLR [15]	0.8411	0.03290
	Proposed Framework	0.9830	0.03958

Table 6. Performance (train/test C-index and IBS) of baseline models using feature selection (feature subsets) selected by LASSO, RSF, and their Union (LASSO+RSF), these methods improve input quality for the proposed GCE framework and address data sparsity.

Feature Set	Model	Train C-index	Test C-index	IBS
LASSO	CoxPH	0.311968	0.307459	0.082238
LASSO	RSF	0.702466	0.678802	0.084416
LASSO	DeepSurv	0.696201	0.669805	0.074979
LASSO	MTLR	0.676988	0.615818	0.080173
RSF	CoxPH	0.278054	0.251886	0.076725
RSF	RSF	0.737036	0.727185	0.077034
RSF	DeepSurv	0.846882	0.828495	0.043584
RSF	MTLR	0.763219	0.786194	0.044982
Union	CoxPH	0.238360	0.215671	0.066030
Union	RSF	0.766275	0.744114	0.065525
Union	DeepSurv	0.896908	0.856182	0.047134
Union	MTLR	0.764552	0.705477	0.051882



(**a**) (**b**)

Figure 3. Comparative evaluation of four survival models (CoxPH, RSF, DeepSurv, MTLR) across three feature selection strategies (LASSO, RSF, and their union) on the SEER cardiac sarcoma cohort: (**a**) C-index (higher is better for discrimination); (**b**)Integrated Brier Score (IBS, lower is better for calibration).

 doi:10.20944/preprints202606.0294.v1

4.3. Performance of Classical and Baseline Models

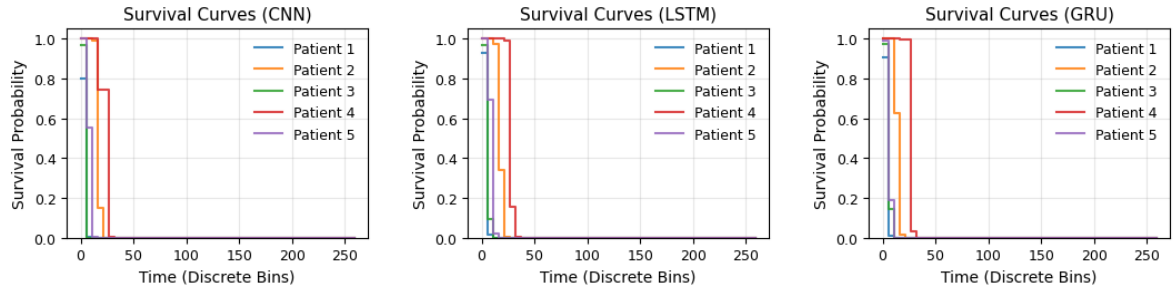
We evaluated four classical and semi-parametric survival models on the 27 selected features without further dimensionality reduction or targeted selection to establish a baseline and contextualize the proposed GCE framework. Table 5 reports their train/test C-index and IBS. DeepSurv performed best among baselines, achieving a test C-index of 0.9061 and the lowest IBS of 0.01990. CoxPH showed solid results (test C-index 0.8842, IBS 0.03280), while MTLR and RSF lagged (test C-index 0.8411 and 0.8322, IBS 0.03290 and 0.04830), indicating reduced effectiveness for capturing complex patterns in this dataset.

4.4. Deep Learning Models and Ensemble Behavior

The DL components and the proposed framework, GCE, were evaluated to assess their ability to integrate nonlinear patterns in the SEER survival data. In Table 3, the GRU model achieved the best single-fold performance among standalone deep models, while the proposed GCE framework further improved results through complementary weighting of GRU and CoxPH. Figure 4 illustrates predicted survival probabilities across time-discretized bins (subfigures a: CNN, b: LSTM, c: GRU).

Training dynamics and generalization are visualized in Figure 5 (training/validation loss over epochs) and (average training and validation loss across folds). All deep models converged stably, with the smallest validation gap, and the ensemble achieved the tightest train–validation alignment, and minimal overfitting in this small cohort. The mean performance across all 10-fold CV, fold-independent summary is provided in Table 3. The standalone GRU model achieved the highest mean C-index of 0.9345, demonstrating exceptional discriminative power.

Kaplan–Meier survival curves are presented in Figure 5 stratified by predicted risk groups (low, medium, high) for the train and test cohort (Figure 5c,d). Low-risk patients maintain substantially higher survival probabilities beyond 100 months. These curves validate the framework's clinical utility: the GCE framework generates well-calibrated, risk-stratified survival distributions that closely align with observed outcomes.



(**a**) (**b**) (**c**)

Figure 4. Predicted survival probabilities across time-discretized bins: (**a**) Survival probability curves predicted by the CNN model, illustrating the decline in survival across discrete time bins and highlighting variability in predicted risk. (**b**) LSTM model curves for patients, showing time-dependent survival dynamics and the ability to capture temporal patterns in patient risk. (**c**) The GRU model produced Survival probability curves for patients, reflecting the temporal progression of survival risk and improved modeling dependencies.

 doi:10.20944/preprints202606.0294.v1

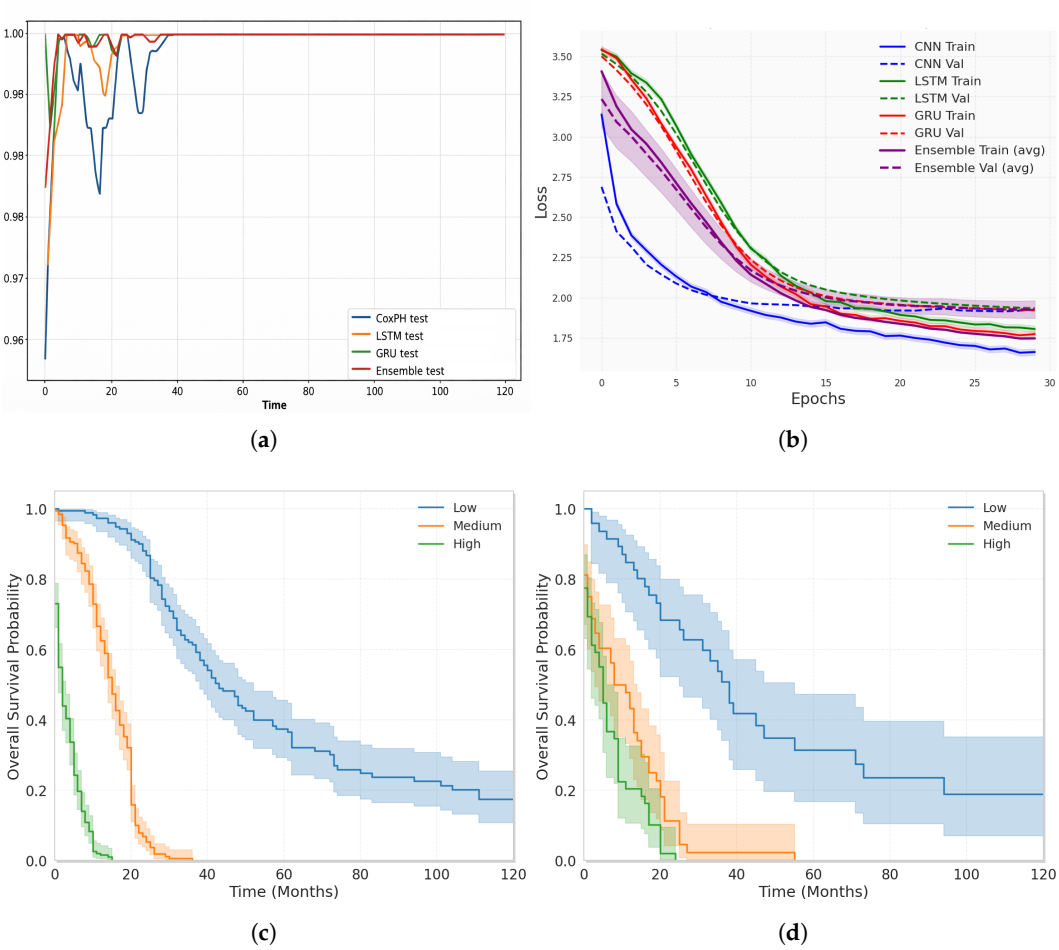


Figure 5. (**a**) Temporal comparison of survival prediction models and Proposed Ensemble Framework achieving the highest mean C-index values over time. (**b**) Proposed Ensemble Framework, Training (solid) and validation (dashed), illustrating stable convergence and improved performance across epochs. (**c**) Kaplan–Meier survival curves for the training cohort, stratified by Low, Medium, and High groups. Low-risk shows the highest survival, while the high-risk indicates a poorer prognosis. (**d**) Kaplan–Meier overall survival curves stratified by low, medium, and high groups, showing the best survival in the low group and the poorest survival in the high group over time.

4.5. Interpretability and Feature Importance Analysis

To ensure clinical trustworthiness and actionable insights into prognostic drivers of primary cardiac sarcoma survival (as shown in Figure ??), SHAP was applied to the proposed GCE framework. The SHAP analysis reveals that our framework relies primarily on clinically meaningful, time-evolving features rather than spurious or redundant signals, enhancing interpretability in a rare-disease context where black-box models are often distrusted. By highlighting the importance of clinically relevant temporal patterns and treatment-related variables, these explanations offer clinicians a transparent rationale for risk stratification and support the framework's superior performance (mean C-index 0.9830, IBS 0.03958) as grounded in biologically relevant drivers rather than overfitting artifacts.

 doi:10.20944/preprints202606.0294.v1

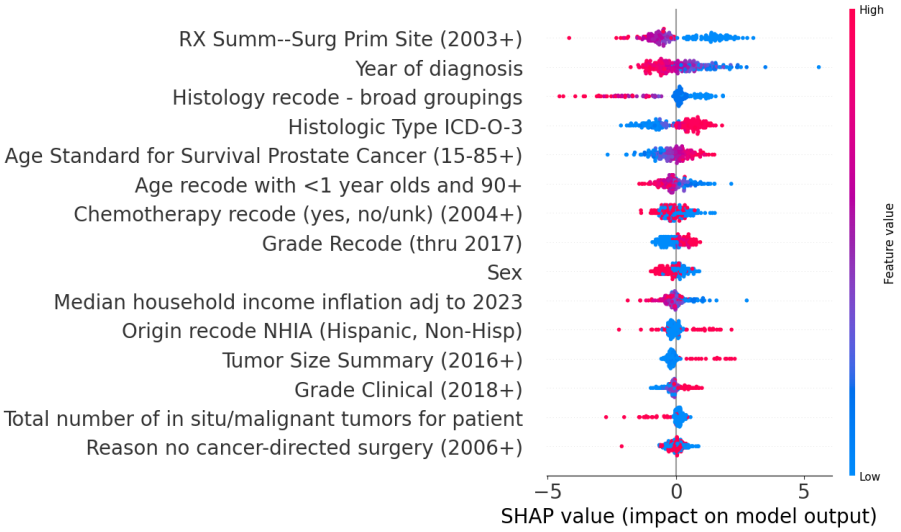


Figure 6. SHAP summary beeswarm plot showing the top predictors of the model, where features are ranked by their contribution, with color indicating low-to-high feature values.

4.6. Comparison with Existing Studies

Our framework substantially outperforms these: mean C-index 0.9830 and IBS 0.03958 exceed prior DL/SEER benchmarks in temporal abstraction. This addresses common gaps in existing works—modest accuracy, overfitting in sparse data, and limited interpretability—while achieving state-of-the-art results in a highly challenging rare-disease setting. The primary discrimination metric Table 7 summarizes key studies.

Table 7. Comparison of recent SEER-based survival prediction studies contextualizes the performance of the proposed GCE framework, compared to recent SEER-based survival prediction studies (2023–2025) employing DL or related ML models. While few directly target cardiac sarcoma (due to its rarity), they focused on similar SEER analyses in other cancers that used recurrent architectures.

Author (Year)	Disease or Dataset	Key Models	C-index (mean/test)	IBS	Our Framework Advantage
[43] (2023)	Gastric adenocarcinoma / SEER	DeepSurv, CoxPH, RSF	0.825–0.871	0.1421	Higher C-index (0.9830); ensemble + SHAP interpretability
[28] (2024)	Hepatocellular carcinoma / SEER	N-MTLR, DeepSurv, RSF	0.824	0.1598	Superior discrimination/calibration; temporal GRU focus
[38] (2025)	Breast cancer / SEER	N-MTLR, CoxPH, RSF	0.771–0.821	0.110	Far higher C-index; addresses small-data engineering gaps
[36] (2024)	Glioblastoma / SEER	XGBoost, DNN, RF	(MSE/RMSE ≈90%)	N/A	Survival-specific metrics (C-index/IBS); fusion robustness
[8] (2023)	Cardiac angiosarcoma / NCDB	Cox regression	(descriptive)	N/A	DL-based prediction; high C-index in related rare tumor
[5] (2023)	Actigraphy Data & Clinical Information	DL models: LSTM, BiLSTM, GRU, RNN	Palliative Performance Index=0.89	N/A	Strong temporal modeling baseline

 doi:10.20944/preprints202606.0294.v1

Table 7. *Cont.*

Author (Year)	Disease or Dataset	Key Models	C-index (mean/test)	IBS	Our Framework Advantage
[32] (2024)	Lung cancer	AI model	C-index = 80.72	N/A	Lower accuracy compared to proposed ensemble
[63] (2025)	Esophageal cancer SEER	CoxPH, RSF, GLMboost, DeepSurv	AUC > 0.81	0.175	Lower calibration in related rare tumor
GCE	**Primary cardiac sarcoma / SEER**	**GCE framework**	**0.9830 (mean)**	**0.03958**	—

5. Discussion

The proposed GCE framework represents a significant advancement in survival prediction for primary cardiac sarcoma to overcome prognostic limitations [6,40,59]. The ensemble achieved a mean C-index of 0.9830 and IBS of 0.03958 across 10-fold CV, outperforming standalone models (e.g., GRU: C-index 0.9345, IBS 0.05105) and classical baselines (e.g., CoxPH: C-index 0.8842, IBS 0.03280). Feature engineering with PCA retaining >90% improved CoxPH/DeepSurv performance (test C-index 0.8719/0.8649), confirming its utility for noise reduction in multidimensional data [22,66]. LASSO/RSF union selection further boosted DeepSurv (C-index 0.856182, IBS 0.047134), demonstrating that ranking preserves prognostic signals better than PCA alone.

The ensemble enhanced calibration, as evidenced by stable loss convergence (Figure 5a,b) and risk-group separation in Kaplan–Meier curves (Figure 5c,d). However, the proposed framework shows a C-index/IBS mismatch: while discrimination is high (0.9830), calibration (IBS 0.03958) is paradoxically worse than CoxPH's simpler IBS (0.03280). This suggests over-confident probability estimates, which we acknowledge as a limitation requiring D-calibration in future work. Additionally, we recognize the absence of competing risks analysis (e.g., Fine-Gray for cardiovascular death), which may inflate C-index estimates in cardiac sarcoma, where cardiotoxicity and heart failure are major competing events.

We also note that the very high C-index (0.983) may reflect the small, homogeneous cohort and requires external validation before clinical deployment. SHAP interpretability revealed a clinically coherent hierarchy, where clinically relevant demographic, tumor-specific, and diagnosis-time temporal covariates contributed most strongly to survival risk prediction, consistent with the disease's aggressive progression [2,3,25]. Crucially, we confirm that no outcome-related variables leaked into the feature set; all SHAP-identified temporal features were derived exclusively from diagnosis-time covariates after rigorous preprocessing, as detailed in Section 3.1.

This post-hoc transparency mitigates DL's black-box criticism, fostering clinical adoption [42]. Building on the framework could incorporate multimodal inputs such as cardiac imaging, which are essential for early tumor detection in cardiac sarcoma [3,60]. Nevertheless, this study lacks external or temporal validation; the reported 10-fold CV results are on a single small SEER cohort (n=727). Prospective or external validation is necessary to confirm generalizability.

From a clinical perspective, the proposed GCE ensemble enables early risk stratification at the time of diagnosis using readily available SEER variables. Patients classified in the high-risk group demonstrated markedly poorer survival, potentially guiding clinicians toward more aggressive therapy [33]. By providing both accurate discrimination and interpretable SHAP-based explanations, the model offers actionable prognostic information. Implications extend to precision oncology, enabling personalized stratification amid cardiotoxicity risks [40,41]. However, retrospective SEER biases (e.g., undercounting sudden deaths) may inflate metrics, warranting caution in extrapolation [26,39].

 doi:10.20944/preprints202606.0294.v1

Limitations

Key limitations include SEER's retrospective biases, incomplete fields (e.g., genetic markers, detailed treatments, comorbidities), and focus on U.S. populations over 65 via SEER-Medicare, limiting generalizability to younger/global cohorts and underestimating rare events like sudden cardiac deaths [39]. The small sample ($n = 727$) and high censoring may contribute to optimistic metrics, despite a 10-fold CV, potentially overlooking external variability [8]. SHAP provides post-hoc interpretability but does not ensure causality, and the absence of multimodal data (e.g., cardiac imaging) restricts diagnosis integration and multimodal modeling [16]. Although outcome-derived variables were removed, retrospective SEER data inherently carry some risk of information leakage; future work will validate the framework on prospective cohorts with strict time-zero feature definitions.

6. Conclusions

This study aimed to improve prognostic prediction for cardiac sarcoma, a rare and aggressive malignancy, by addressing conventional models' inability to capture nonlinear hazard patterns and lack of interpretability. The proposed GCE framework achieved strong performance (mean C-index 0.9830, IBS 0.03958), outperforming standalone DL and traditional models on a small SEER cohort. Key findings demonstrate that a synergistic combination of GRU's temporal learning with CoxPH's statistical anchoring, plus SHAP-based interpretability, enables robust and transparent risk stratification in rare-disease settings. Limitations include retrospective SEER biases and a lack of external validation. Future work requires prospective and multimodal validation (e.g., cardiac imaging) before clinical deployment.

Author Contributions: Muhammad Shoaib Kareem: Writing – original draft, Formal analysis, Data curation, Software, Visualization, Methodology. **Madiha Amjad**: Supervision, Investigation, Formal analysis. **Saba Aslam**: Visualization, Software, Writing – review & editing. **Abdur Rasool**: Supervision, Investigation, Methodology, Validation. **Mutiullah Jamil**: Conceptualization, Validation, Formal analysis. **Hazrat Ali**: Conceptualization, Writing – review & editing.

Funding: Not Given.

Institutional Review Board Statement: Not Applicable.

Informed Consent Statement: Not Applicable.

Data Availability Statement: The dataset (https://seer.cancer.gov/data/) is mentioned in Section III. Code is available at https://github.com/shoaibkareem9-svg/GCE-Framework.git.

Conflicts of Interest: None Declared

References

1. Qiu, Y. L.; Zheng, H.; Devos, A.; Selby, H.; Gevaert, O. A meta-learning approach for genomic survival analysis. *Nat. Commun.* **2020**, *11*, 6350. [CrossRef].
2. Yin, K.; et al. Primary cardiac lymphoma. *J. Thorac. Cardiovasc. Surg.* **2022**, *164*, 573–580. [CrossRef]
3. Bussani, R.; et al. Cardiac tumors: diagnosis, prognosis, and treatment. *Curr. Cardiol. Rep.* **2020**, *22*, 169. [CrossRef]
4. Bangolo, A.; et al. Ten-Year Trends in Hepatocellular Carcinoma Mortality. *Diseases* **2025**, *13*, 256. [CrossRef]
5. Huang, Y.; et al. Deep learning prediction model for patient survival outcomes. *Cancers* **2023**, *15*, 2232. [CrossRtef]
6. Bishnoi, R.; et al. Real-world experience of carfilzomib-associated cardiovascular adverse events. *Cancer Med.* **2021**, *10*, 70–78. [CrossRef]
7. Liao, J.; Zhou, Z. Long-term cardiovascular mortality risk in patients with bladder cancer. *Front. Cardiovasc. Med.* **2023**, *10*, 1142417. [CrossRef]
8. Rahouma, M.; et al. Geographic variation in malignant cardiac tumors and their outcomes. *Front. Oncol.* **2023**, *13*, 1071770. [CrossRef]

 doi:10.20944/preprints202606.0294.v1

9. Cheng, P.; Xie, X.; Knoedler, S.; Mi, B.; Liu, G. Predicting overall survival in chordoma patients using machine learning models. *J. Orthop. Surg. Res.* **2023**, *18*, 652. [CrossRef]
10. Zhang, S.; et al. Personalized prediction for multiple chronic diseases by multi-task Cox learning model. *PLoS Comput. Biol.* **2023**, *19*, e1011396. [CrossRef]
11. Peng, C.; et al. Predicting overall survival in chordoma patients using machine learning models. *J. Orthop. Surg. Res.* **2023**, *18*. [CrossRef]
12. Yan, L.; et al. Deep learning models for predicting survival of chondrosarcoma. *Front. Oncol.* **2022**, *12*. [CrossRef]
13. Zeng, J.J.; Li, K.; Cao, F.; Zheng, Y. Deep learning prognosis prediction of gastrointestinal stromal tumor. *Sci. Rep.* **2024**, *14*. [CrossRef]
14. Liu, Y.; Xie, L.; Wang, D.; Xia, K. Deep learning algorithm for cancer-specific survival in osteosarcoma. *PLoS ONE* **2023**, *18*. [CrossRef]
15. Kiessling, J.; et al. AI outperforms Kaplan–Meier survival estimation. *Eur. J. Vasc. Endovasc. Surg.* **2023**, *65*, 600–607. [CrossRef]
16. Yin, Q.; Chen, W.; Zhang, C.; Wei, Z. CNN model for survival prediction. *Lab. Investig.* **2022**, *102*, 1064–1074. [CrossRef]
17. Tran, K.; et al. Deep learning in cancer diagnosis and prognosis. *Genome Med.* **2021**, *13*. [CrossRef]
18. Vale-Silva, L.A.; Rohr, K. MultiSurv multimodal survival prediction. *medRxiv* **2020**. [CrossRef]
19. Yao, Z.; et al. Multimodal deep learning with imaging and clinical data. *arXiv* **2024**. [CrossRef]
20. Maigari, A.; et al. Multimodal deep learning breast cancer prognosis. *J. Med. Artif. Intell.* **2023**. [CrossRef]
21. Liao, J.; Zhou, Z. Long-term cardiovascular mortality risk in bladder cancer patients. *Front. Cardiovasc. Med.* **2023**, *10*. [CrossRef]
22. National Cancer Institute (NCI). Surveillance, Epidemiology, and End Results (SEER) Program. *SEER Database*. Available online: [CrossRef]
23. Vo, J.B.; et al. Heart disease mortality among breast cancer survivors. *Breast Cancer Res. Treat.* **2022**, *192*, 611–622. [CrossRef]
24. Zheng, Y.; et al. Machine learning in cardio-oncology. *Rev. Cardiovasc. Med.* **2023**, *24*. [CrossRef]
25. Felix, A.S.; et al. Cardiovascular mortality after endometrial cancer. *Int. J. Cancer* **2016**, *140*, 555–564. [CrossRef]
26. Mo, X.; et al. Competing risk analysis in kidney cancer. *BMC Cancer* **2021**, *21*. [CrossRef]
27. Zhang, X.; et al. Predictors of five-year survival in hepatocellular carcinoma. *Cancer Causes Control* **2021**, *32*, 317–325. [CrossRef]
28. Wang, S.C.; et al. Deep learning survival prediction in hepatocellular carcinoma. *Sci. Rep.* **2024**, *14*. [CrossRef]
29. Ezaz, G.; et al. Risk prediction for heart failure after trastuzumab therapy. *J. Am. Heart Assoc.* **2014**, *3*. [CrossRef]
30. Luo, Y.; et al. Heart-specific death in breast cancer patients. *Sci. Rep.* **2025**, *15*. [CrossRef]
31. Xing, H.; et al. Cardiovascular mortality in lung carcinoid tumors. *Medicine* **2023**, *102*. [CrossRef]
32. Caruso, C.M.; et al. Deep learning survival prediction in lung cancer. *Comput. Methods Programs Biomed.* **2024**, *254*, 108308. [CrossRef]
33. Xiao, M.; et al. Survival outcomes of primary cardiac lymphoma. *Hematol. Oncol.* **2020**, *38*, 334–343. [CrossRef]
34. Bishnoi, R.; et al. Carfilzomib cardiovascular events. *Cancer Med.* **2020**, *10*, 70–78. [CrossRef]
35. Hammami, M.B.; et al. Survival outcomes of primary cardiac sarcoma. *Anatol. J. Cardiol.* **2021**, *25*, 104. [CrossRef]
36. Babaei Rikan, S.; et al. Survival prediction of glioblastoma. *Sci. Rep.* **2024**, *14*, 2371. [CrossRef]
37. Zhang, J.; et al. Deep-learning survival prediction of melanoma. *Discov. Oncol.* **2023**, *14*. [CrossRef]
38. Xu, Y.; et al. Deep learning survival prediction in breast cancer. *Sci. Rep.* **2025**, *15*.
39. Faghiri, F.; Kohansal, A. Cox model with Bayesian neural network. *Sci. Rep.* **2025**, *15*. [CrossRef]
40. Al-Badawi, I.A.; et al. Cardiovascular mortality in ovarian cancer. *Medicina* **2023**, *59*, 1476. [CrossRef]
41. Huang, J.; et al. Cardiovascular mortality in Merkel cell carcinoma. *BMC Geriatr.* **2024**, *24*. [CrossRef]
42. Vale-Silva, L.A.; Rohr, K. Multimodal survival prediction. *Sci. Rep.* **2021**, *11*. [CrossRef]
43. Zeng, J.; et al. Survival prediction in gastric adenocarcinoma. *Front. Oncol.* **2023**, *13*, 1131859. [CrossRef]
44. Cao, G.; et al. Postoperative survival prediction for hepatocellular carcinoma. *Res. Square* **2023**. [CrossRef]
45. Sedighi-Maman, Z.; Heath, J.J. Interpretable lung cancer survivability model. *Sensors* **2022**, *22*, 6783. [CrossRef]

 doi:10.20944/preprints202606.0294.v1

46. Pickett, K.; et al. Random survival forests for dynamic predictions. *BMC Med. Res. Methodol.* **2021**, *21*. [CrossRef]
47. Utkin, L.V.; et al. Weighted random survival forest. *Knowl.-Based Syst.* **2019**, *177*, 136–144. [CrossRef]
48. Miandoab, P.; et al. CNN-GRU model for liver tumor tracking. *Med. Phys.* **2025**, *52*. [CrossRef]
49. Asghar, N.; et al. Hybrid CoxPH and DeepHit survival prediction. *BMC Med. Inform. Decis. Mak.* **2024**, *24*, 120. [CrossRef]
50. Tian, D.; et al. ML prognostic model after lung transplantation. *JAMA Netw. Open* **2023**, *6*. [CrossRef]
51. Cai, M.; et al. Random survival forests for spinal chordomas. *J. Clin. Neurosci.* **2025**, *142*, 111697. [CrossRef]
52. Lin, W.; et al. DeepSurv for tongue cancer survival prediction. *J. Craniomaxillofac. Surg.* **2025**, *53*, 1334–1343. [CrossRef]
53. Obite, C.P.; et al. Factor-enhanced DeepSurv model. *Comput. Biol. Med.* **2025**, *189*, 109963. [CrossRef]
54. Aslan, M.F.; et al. CNN-based survival prediction for heart failure. *Biomed. Signal Process. Control* **2021**, *68*, 102716. [CrossRef]
55. Mustafa, E.; et al. Ensemble framework for breast cancer survivability. *Diagnostics* **2023**, *13*, 1688. [CrossRef]
56. Absar, N.; et al. LSTM model for disease outbreak prediction. *Infect. Dis. Model.* **2022**, *7*, 170–183. [CrossRef]
57. Saha, A.; et al. GRU model for COVID-19 patient representation. *ACM BCB* **2023**. [CrossRef]
58. Pradeepa, M.; et al. EfficientNet-GRU model for breast cancer detection. *Sci. Rep.* **2025**. [CrossRef]
59. Zhang, S.; et al. Multi-task Cox model for chronic diseases. *PLoS Comput. Biol.* **2023**, *19*. [CrossRef]
60. Pruitt, S.L.; et al. Survival of pancreatic cancer patients. *Cancer Med.* **2023**, *12*, 200–212. [CrossRef]
61. Zhao, Y.; et al. Wavelet deep learning for cancer prognosis. *BMC Bioinform.* **2020**, *21*. [CrossRef]
62. Katzman, J.; et al. DeepSurv recommendation system. *BMC Med. Res. Methodol.* **2018**, *18*. [CrossRef]
63. Tuersun, A.; et al. Interpretable ML for esophageal cancer survival. *Front. Physiol.* **2025**, *16*. [CrossRef]
64. Yang, X.; Qiu, H.; Wang, L.; Wang, X. Predicting colorectal cancer survival using ML. *J. Med. Internet Res.* **2023**, *25*. [CrossRef]
65. Rasool, A.; Tao, R.; et al. Statistic solution for machine learning to analyze heart disease data. *Proceedings of the 12th International Conference on Machine Learning and Computing (ICMLC)* **2020**, 134–139. [CrossRef]
66. Adam N, Wieder R. Predictive Modeling of Long-Term Survivors with Stage IV Breast Cancer Using the SEER-Medicare Dataset. Cancers (Basel). 2024 Dec 1;16(23):4033. PMID: 39682219; PMCID: PMC11640017. [CrossRef]