Uncertainty-Calibrated Interpretable Tabular Transformer Model for Atrial Fibrillation Prediction with Competing Risk

Jiahui Zhang[1], Aoqi Xie[1], Yiming (Emmett) Peng[1], Yu Shi[1]

[1]Division of Biostatistics, Dalla Lana School of Public Health, University of Toronto

Introduction

Background

Atrial fibrillation (AF) is the most common arrhythmia, affecting up to one-third of individuals over 45. Making timely and accurate AF risk prediction is therefore of significant clinical importance. However, several challenges remain.

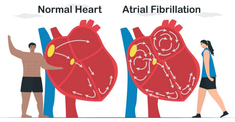


Image Credit: https://www.singhealth.com.sg/patient-care/conditions-treatments/atrial-fibrillation

Challenges and Innovation

Presence of missing values and mixed data types (categorical and continuous)	Use a tabular-specific foundation model
Overlooking complex interactions among features	Leverage a tabular-based foundation model
Lack of uncertainty quantification in risk prediction	Apply conformalized survival analysis for calibrated lower bound survival time
Lack of interpretability in model predictions	Use SHAP for feature-level interpretability

Objectives

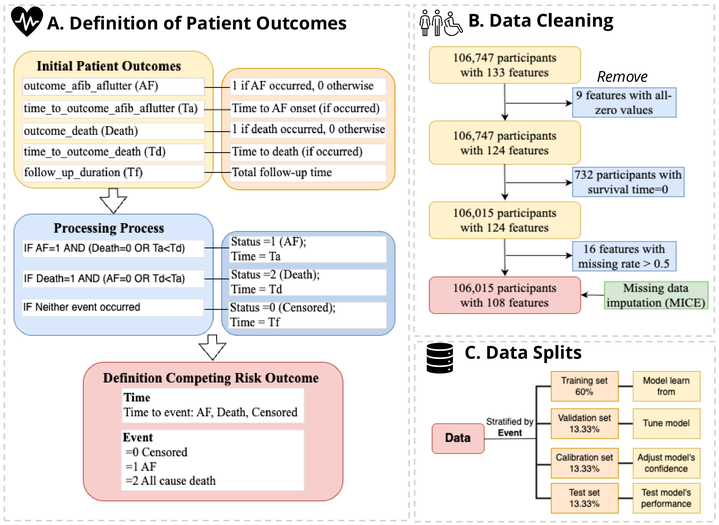
To develop a clinically applicable **competing risk prediction model** that accurately forecasts the **future onset of atrial fibrillation** in individual patients by leveraging a **transformer-based deep survival architecture** to learn contextual embeddings that capture complex feature interactions.

Data

Study cohort

- Randomly selected subset (106,747 patients) from the Cardiovascular Imaging Registry of Calgary (CIROC) with no prior history of AF.
- Patients with a baseline ECG performed between January 2010 and January 2023.

Data pre-processing (Figure 1)



Methodology

Developed an innovative end-to-end model, **UCTT-RP (Uncertainty-Calibrated Tabular Transformer Model for Competing Risk Prediction),** comprising three integrated modules:

Module 1 – Tabular Transformer (Figure 2): Learn robust and informative feature embeddings (low-dimensional representation capturing interactions among features)

- **Input:** $(X_i, \tilde{T}_i, \delta_i)$, where $X_i, \tilde{T}_i, \delta_i$ are features, observed time, and event indicator for individual i, respectively.
- **Output:** $(E_i, \tilde{T}_i, \delta_i)$, where E_i is learned embedding vector for individual i.
- **Design Highlights:** *i. Column embedding* enables the model to distinguish feature identities and handle missing categorical values; *ii. Concatenation* integrates categorical and continuous features to manage mixed data types; *iii. Transformer layers* capture high-order interactions among features.

Module 2 – Deep Survival Models under Competing Risk (Figure 2): Predict the risk of AF occurring at each discrete time point in the presence of competing risk

- **Input:** $(E_i, \tilde{T}_i, \delta_i)$; **Output:** $y_{\tilde{T}_i,\delta_i} = \hat{P}(\tilde{T}_i, \delta_i | E_i)$
- **Model Training:** Fit the survival model (DeepHit) on the training set (D_{train}) to estimate the conditional distribution $\mathrm{F}(\tilde{T}, \delta | E = e)$, $e \in E$
- **Loss Function:** $L_{Total} = L_1 + L_2$. L_1 is the log-likelihood of the joint distribution of the first hitting time and event. L_2 is cause-specific ranking loss function: $\widehat{F_{\delta_i}}(\widetilde{T_i}|e_i) > \widehat{F_{\delta_i}}(\widetilde{T_i}|e_j)$

Module 3 - Conformalized Survival Analysis (Figure 3): Predict lower bound (LB) of survival time

- **Input:** $(X_i, \tilde{T}_i, \delta_i)$
- **Non-conformity score:** Using given trained model, compute $v_{LB}(x, y) = \alpha - F(y)$ on the calibration set (D_{cal})
- **Threshold Selection:** $\hat{q}_{1-\alpha} = 1 - \alpha \frac{|D_{cal}+1|}{|D_{cal}|}$ th quantile of $\{v_{LB}(X_i, \tilde{T}_i)\}$, where α is the target miscoverage rate
- **New Prediction:** $\widehat{LB} = inf\{t \in \mathbb{R}, v_{LB}(x, t) \leq \hat{q}_{1-\alpha}, x \in X\}$ on the test set (D_{test})

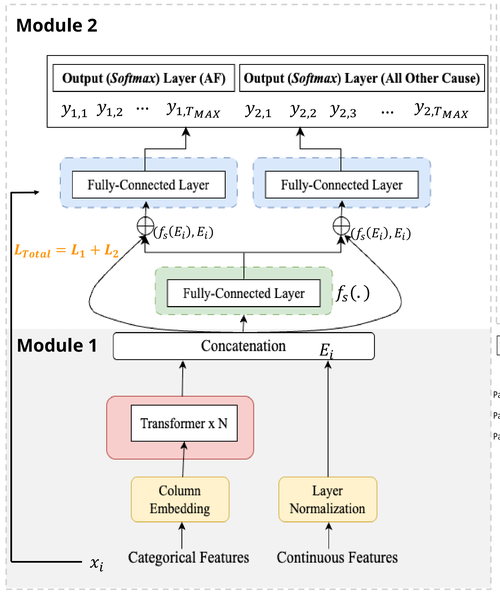


Figure 2: The architecture of the proposed UCTT-RP.

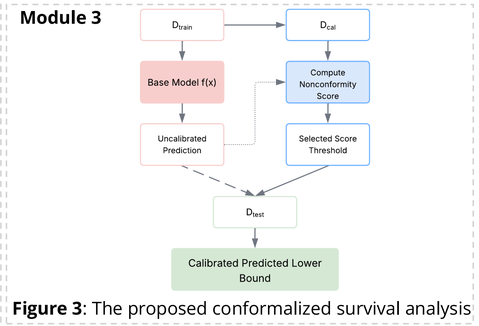


Figure 3: The proposed conformalized survival analysis

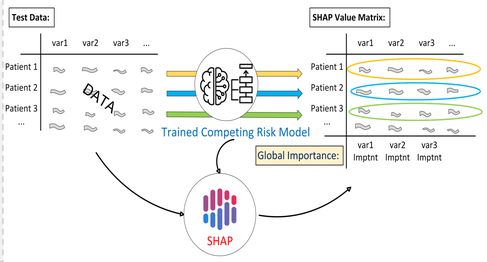


Figure 4: SHAP-based interpretation of AF risk prediction

Interpretability - SHapley Additive exPlanations (SHAP) (Figure 4)

- **Input**: trained model $f(\cdot)$ and $(X_i, \tilde{T}_i, \delta_i)$, where $X_i \in R^p$ and p is the number of input features.
- **Individual and Feature-Specific SHAP Value**: compute $\phi_{ij} = \sum_{S \subseteq F \setminus \{j\}} \frac{|S|!(p-|S|-1)!}{p!} \left[f_{S \cup \{j\}}(x_i) - f_S(x_i) \right]$, where F is the set of all features and S is a subset of features excluding j.
- **Feature-Specific Global Importance**: $\mathrm{Importance}(j) = \frac{1}{n} \sum_{i=1}^{n} |\phi_{ij}|$, which are used to rank features in the SHAP summary plot.

Results

- The proposed model outperforms all benchmark models using C-index as performance metric (**Table 1**), based on results from **500 bootstrap replicates**.
- The proposed model UCTT-RP (End to End) achieves near-exact marginal coverage, other models fail to achieve the target coverage. (**Figure 5**)
- SHAP summary plot (**Figure 6**) shows the top 10 features influencing AF risk. Lipid-lowering therapy use, Age, and random glucose level were the most impactful. Elevated alkaline phosphatase and hematocrit were also associated with increased risk of AF, suggesting two plausible but less established associations.

Table 1: Comparison of Model Performance Using Mean and 95% Confidence Intervals (CI) Across 500 Bootstrap Replicates

Base Model	C-index (95% CI)	LB Coverage (95%CI)
UCTT-RP (End-to-End)	**0.83 [0.825, 0.826]**	0.95 [0.951, 0.952]
UCTT-RP (Two-stage model)	0.81 [0.809, 0.811]	0.93 [0.933, 0.934]
DeepHit (raw features)	0.80 [0.802, 0.803]	0.94 [0.942, 0.943]
Random Survival Forest	0.77 [0.765, 0.766]	0.94 [0.937, 0.938]

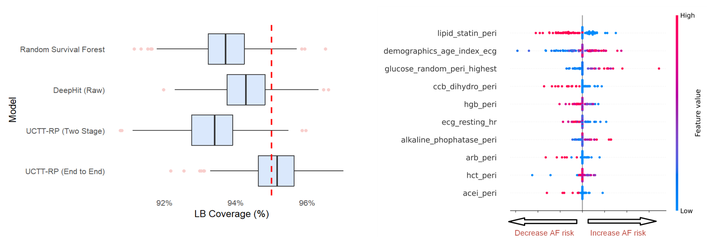


Figure 5: Coverage of each model with target coverage rate at 95%

Figure 6: Individual and feature-specific SHAP value (Order by global importance)

Conclusion

- The proposed end-to-end competing risks model (UCTT-RP) for atrial fibrillation demonstrates strong predictive performance and clinical relevance, with potential applicability to other diseases.
- Conformal prediction enables uncertainty quantification with formal guarantees, but achieving valid coverage still requires accurate base risk estimation.
- SHAP values enhance model interpretability by quantifying each feature's contribution to risk prediction, offering insights into key clinical drivers of atrial fibrillation risk.

Limitations and Future work

- Incorporate other data modalities such as imaging, genomic features, or clinical notes using modality-specific encoders alongside the tabular Transformer backbone.

References and Acknowledgement

1. Emmanuel Candès, Lihua Lei, Zhimei Ren, Conformalized survival analysis, Journal of the Royal Statistical Society Series B: Statistical Methodology, Volume 85, Issue 1, February 2023, Pages 24–45, https://doi.org/10.1093/jrsssb/qkac004
2. Lee, C., Zame, W., Yoon, J., & van der Schaar, M. (2018). DeepHit: A Deep Learning Approach to Survival Analysis With Competing Risks. Proceedings of the AAAI Conference on Artificial Intelligence, 32(1).
3. Huang, X., Khetan, A., Cvitkovic, M., & Karnin, Z. (2020). TabTransformer: Tabular Data Modeling Using Contextual Embeddings. https://doi.org/10.48550/arXiv.2012.06678
4. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Annual Conference on Neural Information Processing Systems (NeurIPS), 30.

We gratefully acknowledge the invaluable support and guidance from our mentor Dr. Pingzhao Hu