Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

ML driven approach for Post-Transplant Surveillance and Risk Modelling in Heart Patients

Dr.K.LATHA Dr.K.LATHA,Assistant Professor(Sl.Grade)

drklatha@aubit.edu.in

University College of Engineering,BIT Campus,Anna University, Tiruchirappalli,Tamilnadu,India.

Article

Keywords: Extreme Gradient Boosting, Random Survival Forest (RSF), Post-Transplant Surveillance, Machine Learning, Hybrid-XR Transplant System

Posted Date: March 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-8391856/v1

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

Additional Declarations: No competing interests reported.

ML driven approach for Post-Transplant Surveillance and Risk Modelling in Heart Patients

Dr.K.LATHA[1*] M.E.,Ph.D.

[1]*Assistant Professor(Sl.Grade),Department of Computer Science and Engineering,

University College of Engineering,BIT Campus,Anna University,

Tiruchirappalli,Tamilnadu,India.

[1]*Email.ID:drklatha@aubit.edu.in,**Orcid ID:**0000-0002-0942-5983

[1]*corresponding author

Abstract

Heart transplantation is the most effective treatment for patients with end-stage heart failure, yet several ongoing challenges limit its success. These include improving donor-recipient compatibility, evaluating surgical risk accurately, and forecasting long-term outcomes reliably. Traditional clinical methods often focus on surgical logistics and lack data-driven tools to guide decision-making throughout the transplant journey. To overcome these limitations, this study presents the Hybrid-XR Transplant System—a modular, machine learning–based framework combining XGBoost and Random Survival Forest (RSF) algorithms.

The system addresses four key areas: donor-recipient matching, pre-operative risk assessment, post-transplant monitoring, and survival prediction. XGBoost is used for classification tasks such as compatibility and risk level, while RSF is applied to survival analysis, offering improved accuracy over traditional statistical models. Each component was rigorously validated using diverse evaluation metrics. Classification performance was measured using Accuracy, Precision, Recall, F1-Score, AUC-ROC, and Matthews Correlation Coefficient (MCC). For survival analysis, the Concordance Index (C-Index) and Log-Rank Test were used.

Hybrid-XR consistently outperformed Support Vector Machines, Logistic Regression, Decision Trees, and the Cox model. By integrating XGBoost-based classification and Random Survival Forest survival analysis, it enables robust interpretation of complex clinical data and supports personalized, data-driven decision-making in heart transplantation.

Keywords: Extreme Gradient Boosting, Random Survival Forest (RSF), Post-Transplant Surveillance, Machine Learning, Hybrid-XR Transplant System.

1. Introduction

1.1. Background

Heart transplantation has long been recognized as a life-saving option for patients suffering from advanced heart failure. While surgical(Kowadlo,2024) techniques have become highly refined, many challenges remain in ensuring successful outcomes. Critical decisions such as selecting suitable donors, evaluating the recipient's risk level before surgery, and managing the patient's condition after the transplant require careful analysis. In most cases, these decisions are based on limited guidelines and the physician's experience, which can vary from case to case. Furthermore, current systems often focus on specific phases of the transplant process and lack a comprehensive approach that supports clinical (Coutance,2022) decision-making throughout the entire patient journey.

This research is driven by the need for a smarter and more integrated solution. With the growing availability of electronic health records and advancements in data science, machine learning presents a powerful opportunity to enhance transplant care. However, existing tools are often fragmented and underutilized, addressing individual problems without offering full-cycle support. The proposed Hybrid-XR Transplant System aims to fill this gap by providing a unified, data-driven framework that supports key stages of the transplant process — from matching donors(Kirk,2020)(Placona,2020;Venturini,2024) and recipients to predicting survival outcomes. By combining classification and survival analysis methods in a modular

design, the system offers clinicians a reliable tool to improve decision accuracy and ultimately enhance patient care.

1.2. Motivation

Heart transplantation(Colvin, 2022) remains one of the most complex and resource-intensive procedures in modern medicine. However, many decisions throughout the transplant process still rely on empirical judgment and generalized scoring systems that may not account for individual patient differences. The on-going shortage of donor(Kirk,2020) (Placona,2020;Venturini,2024) organs further intensifies the need for accurate donor-recipient matching, where a poor match can lead to increased risk of graft failure (Linse,2023;Ohlsson,2023) or rejection. Post-transplant outcomes are also highly unpredictable, with complications such as organ rejection or infection potentially arising at any stage. In addition, forecasting a patient's survival years after the procedure (Kampaktsis,2023) is critical for effective long-term care planning, yet current models frequently fall short in providing reliable estimates.

These limitations highlight the need for an intelligent, data-driven system that can offer meaningful support across the entire transplant life cycle. Existing machine learning tools in this field are often narrow in scope, focusing on single tasks rather than providing a complete decision-support framework. The proposed system addresses this by integrating machine learning techniques capable of analyzing historical clinical data to generate personalized risk predictions, optimize donor-recipient compatibility, and forecast both short- and long-term outcomes. This comprehensive approach is designed to empower clinicians with timely, accurate insights and ultimately improve patient care across every phase of heart transplantation.

1.3. Role of Machine Learning in Clinical Decision Support

Machine learning (ML) has emerged as a trans-formative tool in healthcare, particularly in clinical decision support systems (CDSS). These systems aim to assist medical professionals

by offering data-driven insights that can improve the accuracy, efficiency, and consistency of clinical decisions. Traditional decision-making in medicine often depends on static rules, human expertise, and generalized guidelines, which may not fully capture the complexities of individual patient cases. In contrast, ML models can learn patterns from vast amounts of historical patient data, enabling them to identify hidden correlations, predict risks, and suggest personalized treatment pathways with greater precision.

In the context of heart transplantation(Colvin ,2022), ML holds particular promise due to the highly variable nature of patient responses and the multitude of factors influencing outcomes. By applying algorithms such as classification models and survival analysis techniques, ML can support clinicians in tasks like donor(Kirk,2020) (Placona,2020;Venturini,2024) -recipient matching, pre-transplant risk evaluation, early detection of post-operative(Zafar, 2021;Liu,2023) complications, and long-term survival prediction(Guijo,2020). These insights can help reduce uncertainty, optimize the use of scarce donor organs, and ultimately contribute to more favourable patient outcomes. Integrating ML into clinical workflows also encourages evidence-based medicine, ensuring that decisions are guided not only by experience but also by the best available data.

1.4. System Overview

This Research introduces a multi-module Clinical Decision Support System built upon a comprehensive real-world heart transplantation data-set. The system is organized into four interconnected components, each designed to support a critical stage of the transplant care continuum. The **Compatibility Assessment Module** utilizes machine learning algorithms to determine donor-recipient suitability by analyzing immunological markers, clinical attributes, and demographic data. The **Patient Monitoring Module** employs survival analysis methods to continuously evaluate post-transplant health and identify early signs of complications. The **Outcome Prediction Module** focuses on estimating medium and long-term survival probabilities, aiding clinicians in long-term care planning and resource distribution. Finally,

the **Risk Evaluation Module** categorizes patients based on their risk of adverse outcomes, promoting more individualized and proactive management strategies. Each module is independently trained and validated using rigorous evaluation techniques to ensure clinical accuracy, interpret-ability, and smooth integration into healthcare workflows.

1.5. Objectives

The primary aim of this research is to develop a data-driven Clinical Decision Support System (CDSS) tailored for heart transplantation, with the goal of enhancing clinical decision-making throughout the entire transplant life cycle. By integrating machine learning techniques, the proposed system seeks to overcome the limitations of conventional practices and provide clinicians with timely, personalized, and evidence-based recommendations.

To realize this overarching objective, the research focuses on the following key goals:

- **To improve donor-recipient compatibility** by applying machine learning classification models that consider immunological, clinical, and demographic factors, thereby reducing the risk of early graft rejection(Swamy ,2023;Vaidya,2023) and increasing transplant success rates.
- **To predict short-term survival outcomes** following transplantation using survival analysis techniques, enabling early detection of high-risk patients and facilitating timely clinical interventions.
- **To forecast long-term survival probabilities** over defined intervals (3, 5, and 10 years) by incorporating both static and time-varying patient data, supporting long-term care strategies and efficient resource planning.
- **To perform comprehensive risk stratification** that categorizes patients based on the likelihood of adverse outcomes, allowing clinicians to prioritize cases effectively and adopt proactive treatment approaches.

Each of these objectives contributes to building a robust, interpret-able, and modular framework that aligns with the evolving demands of personalized and predictive transplant medicine.

1.6. Scope of the Project

This research is centered on the design, development, and validation of a comprehensive Clinical Decision Support System (CDSS) that addresses the full spectrum of the heart transplantation(Raghu,2023) process. The system is structured into four intelligent modules, each targeting a specific stage of the transplant life cycle to provide clinicians with actionable insights that enhance decision-making and patient care.

The **Compatibility Assessment** module is responsible for analyzing donor(Kirk,2020) (Placona,2020;Venturini,2024)-recipient matching criteria, aiming to reduce the likelihood of post-operative(Zafar,2021;Liu,2023) complications by ensuring optimal compatibility. The **Patient Monitoring** module uses survival analysis techniques to continuously assess the condition of transplant recipients, identifying patients who may be at increased clinical risk over time. The **Outcome Prediction** module generates probabilistic survival forecasts across medium- and long-term intervals, aiding in strategic care planning. Meanwhile, the **Risk Evaluation** module categorizes patients based on their potential for adverse outcomes, enabling clinicians to prioritize treatment and allocate resources more efficiently.

The development of these models includes advanced methods such as hyper-parameter optimization and survival curve modelling. Visualization tools like Kaplan-Meier survival plots and statistical techniques, including log-rank tests, are applied to interpret and validate model performance, ensuring both clinical relevance and transparency. Overall, the project scope encompasses a modular, data-driven system capable of improving transplant care through intelligent automation and personalized risk analysis.

1.7. Significance of the Research

Integrating machine learning into the heart transplantation process holds significant potential to enhance clinical care through precise, data-driven insights. One of the most impactful benefits is the system's ability to improve donor-recipient matching by leveraging compatibility analysis based on multiple clinical and demographic factors. This approach helps reduce the risk of graft rejection(Swamy ,2023;Vaidya,2023) and increases the likelihood of successful transplant outcomes.

Moreover, the system's predictive capabilities enable early detection of post-transplant complications, allowing healthcare providers to intervene before conditions worsen. By offering reliable long-term survival forecasts, the framework supports strategic planning in post-operative care, resource distribution, and patient monitoring. Taken together, this work represents an important advancement toward a more personalized and proactive model of transplant medicine—where decisions are not just reactive but guided by intelligent, evidence-based tools.

2. Literature Survey: Advancing Heart Transplantation Through Machine Learning and Predictive Analytics

The literature highlights the growing influence of machine learning in transforming heart transplantation through more accurate and data-driven clinical decisions. (Venturini et al., 2024) developed a novel model for predicting 1-year (Zhouet al.,2021) mortality (Peng,2022;Ong,2023;Sinha.2023;Weiss,2023) in paediatrics(Williams,2022) heart transplant patients by simulating hypothetical donor-recipient matches, mortality (Peng,2022;Ong,2023;Sinha,2023;Weiss,2023) outperforming conventional clinical scoring. Similarly, (Zhou et al., 2021) focused on prediction in a Chinese transplant cohort using AI, demonstrating the model's relevance in localized clinical settings. (Agasthi, 2020) and Kilic et al. (2020) emphasized the predictive power of machine learning in identifying long-term graft failure(Linse,2023;Ohlsson,2023) and mortality using comprehensive clinical features,

which outclassed traditional tools. These contributions collectively underscore the potential of machine learning to improve donor-recipient compatibility and support precision-based transplant allocation.

In terms of post-operative care, several studies emphasize the role of machine learning in early complication detection and dynamic patient monitoring. (Colvin et al., 2022) provided a statistical foundation by analyzing national transplant data, while(Linse et al., 2023) and Ohlsson et al. (2023) introduced predictive models targeting early graft (Agasthi, 2020;Mori 2021;Nair 2021;Prabahran,2021) dysfunction within 30 days post-transplant. (Ayers, 2021) and (Aleksova et al.,2020) further supported survival analysis using large-scale datasets and meta-analyses, highlighting the strengths and limitations of current approaches. Notably, (Connor, 2021) and (Gotlieb et al., 2022) addressed the interpret-ability and ethical challenges involved in AI deployment, advocating for models that can support clinicians rather than operate as black boxes. These studies reflect the critical need for real-time, explainable, and clinically validated systems.

Beyond prediction and monitoring, several researchers have expanded machine learning's scope to broader systemic and logistical aspects of transplantation. (Trivedi,2020) and (Slaughter, 2020) and (Vaidya et al., 2023) analyzed the impact of policy reforms on transplant wait-lists and post-operative outcomes, providing insights into how AI models can guide health policy and system design. (Zea-Vera et al., 2023; Peng et al., 2022) explored cost and outcome forecasting after cardiac (Zea,2023)surgery, improving resource planning and hospital management. On the technical frontier, (Gjoreski,2020) demonstrated the use of deep learning(Gjoreski,2020;Zafar, 2021;Yang ,2022;Raghu,2023) for heart failure detection from audio signals, while (Placona,2020) applied natural language processing to donor narratives to optimize matching. Collectively, the literature presents a comprehensive view of how machine learning can be integrated at various points in the heart transplantation life

cycle—from patient selection to long-term outcome prediction—ushering in a new era of personalized and proactive transplant care.

3. Proposed Work

This Research presents the design of an advanced machine learning(Gotlieb, 2022; Miller, 2022;Naruka ,2022;Zea;2022) -based system tailored to support various stages of heart transplantation. The primary aim is to empower clinicians with data-driven tools that can enhance decision-making throughout the transplant process. The proposed framework consists of four key components:

- Donor-Recipient Matching
- Post-Transplant Monitoring
- Long-Term Survival Estimation
- Risk Stratification

Each module is developed to tackle specific clinical challenges by employing customized machine learning(Gotlieb, 2022;Miller, 2022;Naruka ,2022;Zea,2022) models, trained on actual patient data from heart transplant cases. The **Donor-Recipient Matching** and **Risk Stratification** modules both leverage the **XGBoost algorithm**. The first focuses on predicting compatibility between donors and recipients using factors such as blood group, **HLA mismatch**, and **CMV status**. The latter categorizes patients into **low**, **moderate**, or **high-risk** groups to aid in prioritizing care and allocating medical resources efficiently.

The **Post-Transplant Monitoring** and **Survival Prediction** components utilize **Random Survival Forests (RSF)**. These models help analyse complications that may arise after surgery(Kampaktsis,2023) and forecast survival probabilities at various intervals—specifically at 3, 5, and 10 years post-transplant using techniques from survival analysis.

To assess the effectiveness of each component, appropriate performance metrics are employed, including **accuracy**, **ROC-AUC**, and **survival curve analysis**. Overall, the

system is designed to improve the accuracy of diagnoses and prognoses, support personalized patient care, and optimize risk evaluation in the context of heart transplantation.

3.1. Proposed System Workflow Steps: From Data to Decision Support

Step 1: Data Acquisition & Intelligent Pre-processing

Raw donor and recipient clinical data are collected as the initial input. This data undergoes essential pre-processing steps, including missing value handling, encoding of categorical variables, normalization, and feature selection—ensuring the data-set is clean, consistent, and ready for analysis.

Step 2: Smart Donor-Recipient Matching (XGBoost Engine)

An advanced matching mechanism based on the **XGBoost algorithm** evaluates compatibility between donor and recipient pairs. Clinical attributes such as **blood group**, **HLA mismatch**, and **CMV status** are utilized to generate a **match probability score**, guiding optimized donor selection.

Step 3: Post-Transplant Health Monitoring (RSF Model)

Following transplantation, the patient's progress is monitored using a **Random Survival Forest (RSF)** model. This stage produces **dynamic survival probabilities** and **Kaplan-Meier survival curves**, enabling early detection of potential complications and supporting timely clinical interventions.

Step 4: Long-Term Prognosis Forecasting (RSF-Based Survival Estimation)

The RSF model is also leveraged to generate **long-term survival forecasts** at **3, 5, and 10-year** intervals post-transplantation. These insights help in strategic planning for on-going care and personalized patient management.

Step 5: Risk Stratification & Clinical Prioritization (XGBoost Classification)

In the final stage, an **XGBoost classifier** assigns a **risk label** to each patient—**Low**, **Moderate**, or **High**. This classification helps healthcare providers prioritize treatment and optimize the allocation of medical resources based on predicted outcomes.

3.2. Materials, Models and Methodology

3.2.1. Extreme Gradient Boosting (XGBoost) Algorithm

XGBoost, short for **Extreme Gradient Boosting**, is a highly efficient and widely adopted technique used for both classification and regression tasks. It belongs to the family of **boosting algorithms**, where multiple weak learners—typically decision trees—are combined to form a strong predictive model. Unlike traditional models that attempt to capture patterns in one pass, XGBoost constructs decision trees sequentially. Each new tree is designed to correct the mistakes made by the ensemble of previously built trees. This iterative correction process enhances predictive performance by reducing the overall error. What sets XGBoost apart from standard gradient boosting methods is its speed and accuracy. It utilizes a **gradient descent approach** to minimize the loss function and incorporates **second-order derivatives** (via Taylor expansion) for improved optimization. This makes the learning process more precise and stable.

A key feature of XGBoost is its **regularization capability**, incorporating both L1 (Lasso) and L2 (Ridge) penalties. This helps in preventing over-fitting and improving the generalization of the model on unseen data. The algorithm is also capable of **handling missing values automatically**, which makes it robust when dealing with real-world datasets. Another advantage is its **support for parallel processing**, allowing it to train models significantly faster compared to traditional boosting techniques. XGBoost also computes **feature importance scores**, offering interpret-ability by highlighting which variables most influence the predictions.

It supports both **binary and multi-class classification**, and outputs **probability scores** for classification problems. Additionally, the algorithm performs well even in the presence of **noisy data or outliers**, making it a strong choice for complex and large-scale prediction tasks. XGBoost is highly configurable through its various hyper-parameters, such as the

learning rate, **maximum tree depth**, and **number of estimators**, which can be fine-tuned to optimize model performance.

In summary, XGBoost is a **fast, flexible, and accurate** machine learning(Guijo,2020; Zhong,2021)(Connor,2021) algorithm that is widely used across industries and academic research for developing robust predictive models.

- **Donor-Recipient Matching**

The initial component of the proposed system is dedicated to evaluating the **compatibility between donors and recipients** prior to heart transplantation. Ensuring a biologically and immunologically suitable match is crucial to minimizing the risk of graft rejection(Agasthi, 2020;Mori 2021;Nair, 2021;Prabahran,2021) and post-operative complications. To accomplish this, the module employs the **XGBoost algorithm**, a highly efficient gradient boosting technique known for its strong performance in classification tasks. This model is structured as a **binary classifier**, designed to distinguish between compatible and incompatible donor-recipient pairs.

The model is trained on a curated set of key clinical features, including **age, gender, race, BMI, blood type (A, B, O), HLA mismatch count, panel reactive antibody (PRA) levels, CMV match status, and cross-match results (CRSMATCH_DONE)**. These features collectively capture immunological and physiological compatibility indicators. To comprehensively evaluate all potential matches, the data-set is transformed into a **Cartesian product** of donor and recipient records, generating every possible pairing. Each pair is then labeled based on historical compatibility outcomes—either a match or no match. Once trained, the XGBoost model assigns a **probabilistic match score (ranging from 0 to 1)** for each donor-recipient pair. These scores enable ranking and prioritization, helping clinicians identify the most suitable recipients for available donor organs. High-scoring pairs are flagged as optimal candidates for transplantation.

Model performance is rigorously assessed using classification metrics like **accuracy, precision, recall, and the area under the ROC curve (ROC-AUC)**. The use of such metrics ensures that the model not only makes correct predictions but also maintains clinical relevance and reliability. By integrating this predictive module into the organ allocation process, healthcare teams can accelerate decision-making, **reduce wait-list durations**, and enhance overall transplant success rates.

- **Risk Stratification and Clinical Summarization**

This module is designed to categorize heart transplant recipients into clearly defined risk levels—Low, Moderate, or High—based on their individual clinical profiles. Accurate risk stratification is essential for guiding post-transplant care, optimizing resource allocation, and ensuring timely medical intervention for high-risk patients.

To achieve this, a **multi-class classification model** is developed using the **XGBoost algorithm**, known for its speed, accuracy, and robustness in handling structured clinical data. The model leverages a rich array of input features, encompassing patient demographics, laboratory values, and critical post-transplant clinical indicators. During model training, **stratified sampling** is employed to maintain the correct distribution of risk categories across both training and validation datasets. This ensures balanced learning and helps the model generalize effectively across different patient groups. Upon training completion, the model's performance is assessed through standard evaluation metrics like **accuracy, F1-score, precision, recall, and ROC-AUC**, offering a comprehensive view of its predictive ability.

For ease of use in real-world clinical settings, an **interactive input interface** is incorporated. This allows healthcare professionals to input specific patient data and receive instant, interpret-able risk labels—clearly displayed as **"Low Risk," "Moderate Risk," or "High Risk."** This eliminates the need for complex post-hoc explanations and facilitates straightforward decision-making. By automating the classification process, this module empowers clinicians with a fast and reliable tool for **prioritizing patient care**, streamlining

post-operative monitoring, and planning timely interventions for those at elevated risk. Ultimately, it enhances the quality and efficiency of transplant follow-up strategies.

3.2.2. Random Survival Forest (RSF)

The **Random Survival Forest (RSF)** algorithm is an advanced ensemble learning method specifically tailored for survival analysis, where the primary objective is to predict the time until the occurrence of a particular event—such as patient mortality(Ayers ,2021; Prabahran,2021) or system failure(Linse,2023;Ohlsson ,2023). Unlike standard regression or classification models, RSF is designed to work with **time-to-event data**, which often includes **censored observations**—cases where the outcome event has not yet occurred during the observation period. To address this, RSF builds a collection of survival trees using **bootstrap sampling**, and each tree is constructed using a modified splitting criterion that evaluates survival differences between groups using statistical measures like the **log-rank test**. Rather than outputting a fixed value, each tree in the forest produces a **survival function**, representing the estimated probability of survival over time for a given input.

By aggregating the survival curves from all trees in the ensemble, RSF generates a comprehensive prediction that captures the temporal dynamics of the survival process. This approach is **non-parametric**, meaning it does not rely on predefined assumptions about the shape or distribution of survival times, making it highly adaptable to real-world clinical data. RSF is capable of estimating key survival outcomes such as **median survival time**, **hazard functions**, and **cumulative survival probabilities**, while also handling missing values and both categorical and continuous variables with ease. Its utility is particularly evident in medical domains, where accurate risk estimation over time is essential. The algorithm provides **variable importance measures**, which help identify the most influential factors impacting patient outcomes. Model performance is typically assessed using the **Concordance Index (C-index)**, which evaluates the agreement between predicted and observed survival

rankings. Overall, RSF is a robust and interpret-able tool for modelling survival data, especially when traditional methods fall short in handling censored and complex datasets.

- **Post-Transplantation Monitoring**

The **post-transplantation monitoring module** is designed to track the health status of heart transplant recipients over time and estimate their survival probabilities at various intervals. This module applies the **Random Survival Forest (RSF)** algorithm, which is an ensemble learning technique well-suited for survival analysis involving censored data. Censoring occurs when the full outcome (such as death(Khera,2021;Gjesdal, 2022;Ayer,2023) or relapse) is not observed for every patient, either because the event hasn't happened yet or due to loss to follow-up. RSF handles this complexity effectively by building multiple decision trees from bootstrapped samples and using **log-rank statistics** to determine optimal splits at each node. Features such as **creatinine level (CREAT_REC)**, **blood urea nitrogen (BUN_REC)**, **ischemic time (ISCHTIME)**, **left ventricular ejection fraction (LV_EJECT)**, **transfusion history**, and **steroid use** are used as input to model post-transplant outcomes.

Each RSF tree contributes to generating a **personalized survival curve** for the patient, illustrating the probability of survival across time. These curves are aggregated to provide a more accurate survival estimate and a visual representation of patient risk. This component not only predicts survival probability but also aids clinicians in dynamically adjusting post-operative care based on the patient's evolving risk profile. Performance evaluation is done using the **Concordance Index (C-index)**, which measures how well the model's predictions align with actual survival rankings. By integrating survival modelling into routine post-transplant care, this module promotes **early identification of potential complications**, allowing for timely interventions and **individualized follow-up planning**, thereby improving long-term transplant outcomes.

- **Long-Term Survival Forecasting Module**

The third module is dedicated to estimating the **long-term survival probabilities** of heart transplant recipients at critical post-operative intervals—specifically at **3, 5, and 10 years**. Leveraging the **Random Survival Forest (RSF)** technique, this component extends the approach used in post-transplant monitoring by shifting its focus toward **long-term risk estimation**. The model is trained on a refined set of variables that include **ischemic time (ISCHTIME)**, **HLA mismatch**, **maximum panel reactive antibody levels (PRA_MAX)**, **left ventricular ejection fraction (LV_EJECT)**, **steroid usage**, **blood transfusion records**, and outcome indicators derived from survival status at 365 days. These features enable the system to predict how a patient's survival likelihood evolves far beyond the immediate post-operative window.

For each recipient, the RSF model generates **personalized survival curves**, such as **Kaplan-Meier estimates**, illustrating the probability of survival over **1,095 days (3 years)**, **1,825 days (5 years)**, and **3,650 days (10 years)**. In addition to these curves, the module outputs **patient-specific survival scores** and **survival histograms** to visually represent the distribution of survival likelihoods across the cohort. The predictive performance is validated using the **Concordance Index (C-index)**, which quantifies the model's ability to accurately rank survival outcomes. This module plays a vital role in **long-term clinical planning**, enabling physicians to identify patients at risk for **late-onset complications or graft failure**(Agasthi, 2020;Mori 2021;Nair 2021;Prabahran ,2021) and to develop **tailored care strategies** that can extend patient longevity and improve quality of life(Zhong,2021) following transplantation.

4. System Architecture

Fig. 1 illustrates the individual components of the proposed -based framework designed for heart transplantation. Each module plays a specific role, from donor-recipient matching to long-term survival prediction, forming an integrated system aimed at enhancing clinical outcomes.

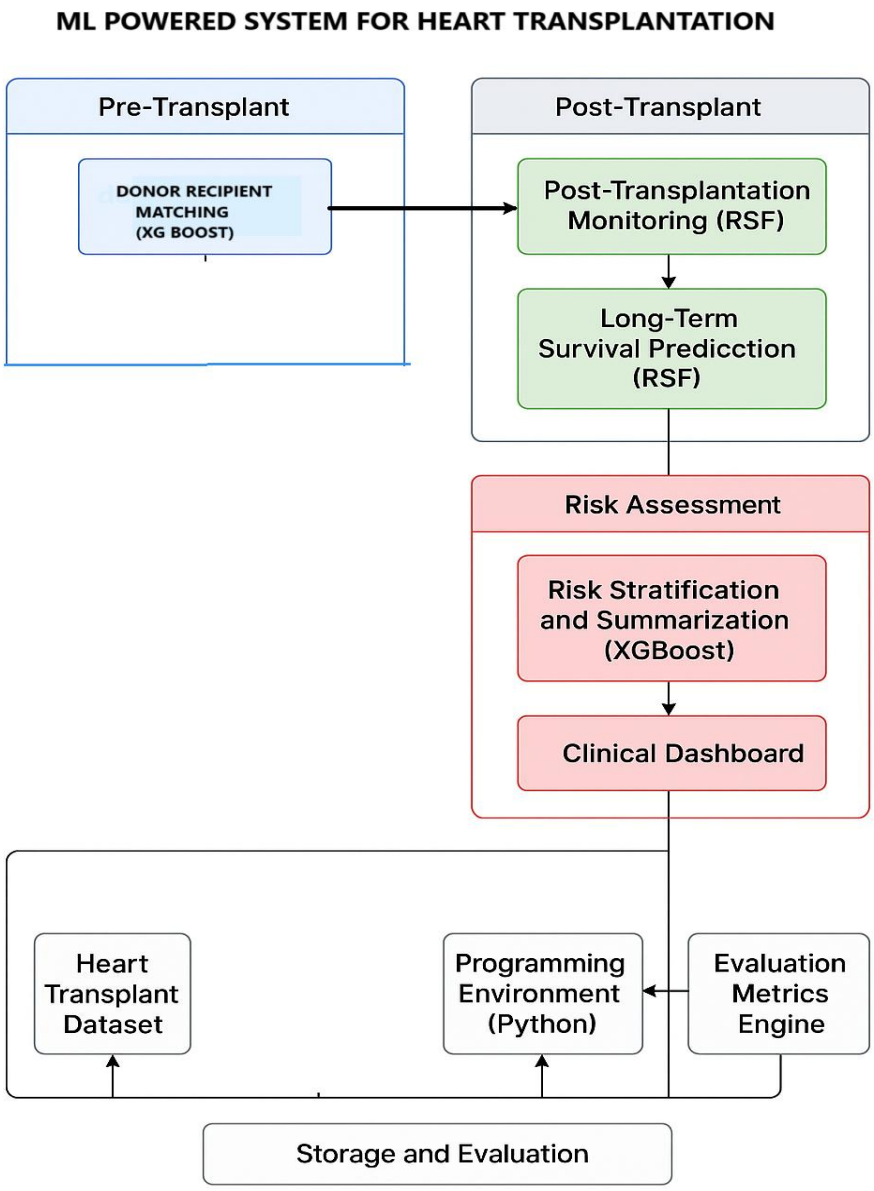


Fig. 1: System Architecture

4.1. Component –Wise Overview of the Proposed ML-Based Heart Transplant Framework

Donor-Recipient Matching

- **Objective:**

To determine optimal compatibility between donors and recipients before transplantation, aiming to reduce the chances of organ rejection and improve surgical outcomes(Kowadlo,2024).

- **Model Implemented:** Extreme Gradient Boosting (XGBoost) – Binary classification model.
- **Input Features:**
    - Donor and recipient characteristics: age, gender, race, body mass index (BMI)
    - Immunological and biological markers: blood type (A, B, O), HLA mismatch levels, CMV status, and CRSMATCH_DONE indicator.
- **Methodology:**

  The input data is transformed into all possible donor-recipient pairings using a Cartesian product. Each pair is labeled as either compatible or incompatible. The XGBoost algorithm is trained on these labeled instances, learning patterns from both clinical and immunological data to predict compatibility.
- **Outputs Generated:**
    - Compatibility score between 0 and 1 for each pair
    - A ranked list of the most suitable matches
    - Evaluation metrics: Accuracy, Precision, Recall, and ROC-AUC for validation

Post-Transplantation Monitoring

- **Objective:**

  To provide continuous assessment of recipient health status following surgery and to predict short-term survival probabilities dynamically.
- **Model Implemented:** Random Survival Forest (RSF) – A survival analysis model.
- **Input Features:**
    - Post-operative indicators: Serum creatinine (CREAT_REC), Blood Urea Nitrogen (BUN_REC), Ischemic Time (ISCHTIME), Left Ventricular Ejection Fraction (LV_EJECT), Number of transfusions, and Cortico-steroid use.

- **Methodology:**

  RSF constructs an ensemble of decision trees, utilizing bootstrapped training sets and log-rank tests at each node for survival outcome separation. It accommodates censored data and clusters patients with similar survival profiles.

- **Outputs Generated:**

- Individualized Kaplan-Meier survival curves
- Time-dependent survival probabilities for each patient
- Early alerts for possible complications
- Model validated using the Concordance Index (C-index)

Long-Term Survival Prediction

- **Objective:**

  To forecast long-range survival probabilities post-transplant at distinct milestones (3, 5, and 10 years), supporting proactive care planning.

- **Model Implemented:** Random Survival Forest (RSF) – Extended for time-based prognosis.

- **Input Features:**

- Clinical predictors: ISCHTIME, HLA mismatch score, maximum PRA level (PRA_MAX), LV_EJECT, corticosteroid usage, and number of transfusions.
- Labelled survival data tailored to specific post-operative durations

- **Methodology:**

  The model focuses on generating long-term predictions by mapping patients' clinical features to expected survival durations. RSF produces probabilistic survival distributions and Kaplan-Meier-type curves for each interval.

- **Outputs Generated:**

- Survival likelihoods at 3, 5, and 10 years (1095, 1825, and 3650 days)
- Graphical outputs: survival histograms, risk timelines

- Evaluation using C-index for performance consistency

Risk Stratification and Clinical Summarization

- **Objective:**

  To classify transplant patients into Low, Moderate, or High-risk categories (Colvin,2022) thereby enhancing treatment prioritization and patient monitoring strategies.

- **Model Implemented:** XGBoost – Multi-class classification model.
- **Input Features:**
  - Diverse parameters including demographic data, lab results, intra and post operative vitals, and donor-recipient compatibility attributes.
- **Methodology:**

  Training and testing datasets are prepared with stratified sampling to preserve the distribution of risk categories. The XGBoost model is trained to distinguish between the three risk levels based on multi-dimensional input features.

- **Outputs Generated:**
  - Risk category assignment for each patient (Low, Moderate, High)
  - Evaluation metrics: Classification accuracy, F1-Score, and Confusion Matrix
  - Integrated prediction interface for real-time clinical application

5. Result and Discussion

5.1. Data-set Description and Clinical Relevance

The dataset used in this study (**cleanedDataset365.csv**) comprises approximately **43,000 records** describing heart transplant cases. For the analysis presented in this work, **39 clinically relevant attributes** were considered for model development and evaluation. These attributes encompass recipient and donor details such as age, sex, blood type, body mass index (BMI), ischemic time, and immunosuppressive therapy indicators. The dataset is well-structured, features minimal missing data, and is highly suitable for machine learning

(Guijo,2020; Zhong,2021;Connor,2021)applications. Its diversity in patient profiles and clinical parameters enables the development of predictive models for various tasks, including donor-recipient compatibility, monitoring post-operative health, and estimating long-term survival. The dataset was used as obtained from the public source. The dataset is publicly available via **Zenodo** and can be accessed at **https://doi.org/10.5281/zenodo.18030387**.

5.2. Comparative Evaluation of Classification and Survival Models

An analytical comparison of classification algorithms reveals critical observations concerning performance metrics like accuracy, precision, recall (sensitivity), specificity, and F1-score. Generally, a higher accuracy aligns with improved precision and specificity, reflecting the model's ability to correctly identify both true positives and true negatives. While there is often a trade-off between precision and recall, well-optimized models strike a functional balance, ensuring both reliability and predictive robustness. The F1-score, a harmonic mean of precision and recall, effectively summarizes this trade-off and serves as a key indicator of a model's comprehensive performance. Models demonstrating consistent results across these metrics tend to generalize better and present a lower risk of over-fitting, thus offering stability when applied to unseen clinical data.

In survival analysis, model performance is assessed using the Concordance Index (C-Index), which measures the model's ability to correctly rank individuals based on their survival probability. A higher C-Index suggests superior discriminatory capability. This index, when paired with a lower Log-Rank statistic—indicating minimal divergence between observed and predicted survival curves—reflects a model with strong generalization and minimal over-fitting. Conversely, a moderate C-Index combined with a high Log-Rank value may signal unreliable prediction behaviour. Therefore, models achieving both a high C-Index and a low Log-Rank score are considered optimal for clinical implementation, offering dependable prognostic value in long-term patient outcome prediction.

Table1: Performance metrics for XGboost

Model	Accuracy	Precision	Sensitivity	Specificity	F1-Score	Over-fitting Risk
SVM	0.92	0.90	0.88	0.94	0.89	0.015
Logistic Regression	0.91	0.89	0.87	0.93	0.88	0.01
Random Forest	0.93	0.91	0.89	0.95	0.90	0.03
Decision Tree	0.89	0.86	0.83	0.90	0.84	0.10
XGBoost	0.95	0.94	0.90	0.96	0.92	0.015

Table 1 presents a comprehensive evaluation of various classification algorithms based on key performance metrics, including Accuracy, Precision, Sensitivity (Recall), Specificity, F1-Score, and Over-fitting Risk. Among the models tested, the XGBoost classifier demonstrated the most robust overall performance, achieving the highest accuracy of 95% and an F1-score of 0.92. These results highlight its effectiveness in maintaining a balance between correctly identifying positive cases and minimizing false classifications.

Moreover, XGBoost exhibited a low over-fitting risk (1.5%), comparable to that of the Support Vector Machine (SVM), but surpassed it in both precision and recall, thus contributing to its superior F1-score. Although the Random Forest model also showed strong results with an accuracy of 93%, its higher over-fitting risk (3%) made it a less optimal choice. In contrast, simpler models such as the Decision Tree recorded lower accuracy (89%) and a significantly elevated over-fitting risk (10%), reducing their suitability for clinical deployment.

Overall, the analysis confirms that XGBoost provides the most reliable trade-off between predictive accuracy, generalizability to unseen data, and model stability. As such, it stands

out as the most appropriate model for implementing risk stratification in heart transplant (Goff,2020) outcome prediction frameworks.

5.3. Evaluation Metrics for Classification Models

To assess the performance of classification algorithms in the heart transplant(Goff,2020) prediction system, several standard evaluation metrics are utilized. Each metric provides insight into different aspects of a model's performance.

1. **Accuracy:** Accuracy represents the ratio of all correctly predicted outcomes—both positive and negative—to the total number of predictions made. It provides an overall effectiveness measure of the model's correctness.

$$Accuracy = TP + TN/TP + TN + FN + FP \quad (1)$$

Where:

- **TP (True Positives)**: Positive cases correctly predicted as positive
- **TN (True Negatives)**: Negative cases correctly predicted as negative
- **FP (False Positives)**: Negative cases incorrectly predicted as positive
- **FN (False Negatives)**: Positive cases incorrectly predicted as negative

2. **Precision:** Precision quantifies how many of the cases predicted as positive are actually positive. It is particularly important when false positives are costly.

$$Precision = \frac{TP}{TP} + FP \quad (2)$$

3. **Sensitivity (Recall or True Positive Rate):** Sensitivity, also known as Recall, measures the ability of the model to correctly identify all actual positive cases.

$$Sensitivity\ (Recall) = TP/TP + FN \quad (3)$$

4. **Specificity (True Negative Rate):** Specificity indicates the proportion of true negatives correctly identified by the model. It is crucial in contexts where false positives must be minimized.

$$\text{Specificity} = \frac{\text{TN}}{\text{TN+FP}} \quad (4)$$

5. **F1-Score:** The F1-Score is the harmonic mean of Precision and Recall. It is particularly useful when the class distribution is imbalanced and a single measure is needed to balance both metrics.

$$F1 = 2 \cdot Precision. \frac{Recall}{Precision} + Recall \quad (5)$$

6. **Over-fitting Risk:** Over-fitting is identified by a significant gap between training and testing accuracy. A model that performs very well on training data but poorly on new data may be over-fitting. A large difference suggests that the model is not generalizing well, and adjustments such as regularization or cross-validation may be necessary.

$$Overfitting = Training\ Accuracy - Testing\ Accuracy \quad (6)$$

5.4. Comparative Evaluation of Survival Models

Table 2 provides a comparative evaluation of different survival models based on two critical metrics: the Concordance Index (C-Index) and the Log-Rank statistic. The C-Index assesses how well the model ranks patients by their survival outcomes, while the Log-Rank test evaluates the statistical difference between survival curves across groups.

Table 2: Performance Metrics for Random Survival Forest

MODEL	C-Index	Log-Rank
Cox Proportional Hazards	0.71	18.2
GBM-Survival	0.74	16.9
Kaplan-Meier	0.68	20.3
Random Survival Forest	0.81	12.6

Among the models examined, the Random Survival Forest (RSF) demonstrated the strongest performance, achieving a C-Index of 0.81, which reflects its high accuracy in ranking survival times. In addition, RSF maintained a relatively low Log-Rank value of 12.7,

indicating minimal divergence between predicted and observed survival trends, thereby suggesting better generalization and lower risk of over-fitting. In contrast, traditional models such as Kaplan-Meier yielded a lower C-Index of 0.68, along with a higher Log-Rank value, highlighting their limited predictive discrimination and weaker curve separation. These findings confirm that RSF offers a balanced and reliable approach for survival prediction, effectively combining accurate prognostic ranking with stable curve estimation across diverse patient profiles.

1. Concordance Index (C-Index)

The Concordance Index, commonly referred to as the C-Index, is a performance metric used to evaluate the discriminative ability of survival models. It quantifies how well the model can rank individuals based on their predicted risk or survival time. The index is calculated as:

$$C = \frac{Number\ of\ concordant\ pairs}{Number\ of\ comparable\ pairs} \quad (7)$$

A *concordant pair* refers to a pair of patients where the individual with the shorter survival time is correctly assigned a higher predicted risk. A *comparable pair* includes any two individuals whose survival times can be meaningfully compared—specifically, cases where the earlier event is not censored. A C-Index value of 0.5 implies random predictions, while a value closer to 1.0 indicates excellent predictive discrimination.

2. Log-Rank Test (χ^2 Statistic): The Log-Rank test is a statistical method used to determine whether there are significant differences between the survival distributions of two or more groups. It tests the null hypothesis that there is no difference in survival experience between the groups. The test statistic is defined as:

$$\chi^2 = \sum_{j=1}^{k} \frac{(O_j - E_j)^2}{E_j} \quad (8)$$

- O_j is the observed number of events (such as deaths) in group *j*,
- E_j is the expected number of events under the null hypothesis,

- k is the number of groups being compared.

A higher chi-square value indicates a greater difference between the groups' survival curves, potentially leading to the rejection(Trivedi,2020) of the null hypothesis. This test is particularly useful for evaluating how well survival models distinguish between patient risk categories over time.

Table 3: Hyper-parameter Settings for XGboost and RSF Components

Components	**Hyperparameter**	**Value range**
XGBoost(Module 1 & 4)	Learning Rate	0.01-0.3(selected:0.1)
	Number of Estimators	100-1000(selected:300)
	Max Depth	3-10(selected: 6)
	Subsample	0.5-1.0(selected: 0.8)
	Gamma	0-5(selected:0.1)
	Column sample by tree	0.5-1.0(selected: 0.7)
	Regularization(Lambda)	Ridge(λ=60) Lasso(λ=100)
Random Survival Forest (Module 2 & 3)	Number of estimators	100-1000(selected: 500)
	Min samples split	2-10(selected: 4)
	Min samples leaf	1-5(selected: 2)

5.5. Hyper-parameter Configuration for XGBoost and RSF Models

Table 3 outlines the specific hyper-parameter configurations applied to the XGBoost and Random Survival Forest (RSF) models used in the study. These settings were carefully selected based on iterative tuning and validation to optimize model performance. For XGBoost, parameters such as learning rate, maximum tree depth, and the number of estimators were adjusted to enhance classification accuracy. Similarly, for the RSF component, key parameters including the number of trees, node size, and splitting rules were fine-tuned to improve survival prediction reliability. This table serves as a reference for the

model setup and re-producibility of results. Further, to improve the model's robustness and minimize variance, sub-sampling and column sampling rates were adjusted to **0.8** and **0.7**, respectively. A **gamma value of 0.1** was used to introduce a threshold for tree splits; ensuring only significant partitions were created. To prevent over-fitting and enhance model generalization, **L2 regularization ($\lambda = 60$)** and **L1 regularization ($\lambda = 100$)** were incorporated.

For the RSF models used in Modules 2 and 3, parameter settings were chosen to align with the demands of survival data. A forest comprising **500 trees** was used to ensure both prediction stability and computational feasibility. The minimum number of samples required to split an internal node was set to **4**, limiting unnecessary branching based on minor variations. Similarly, **a minimum of 2 samples per terminal node** was established to ensure sufficient data in each leaf for credible survival estimation. These RSF configurations significantly improved the **C-Index**, suggesting high accuracy in ranking patient survival probabilities. At the same time, the **Log-Rank statistic** remained low, indicating minimal deviation between predicted and actual survival curves. Together, these hyper-parameter choices enhanced the clinical reliability and deployment-readiness of the model for real-world transplantation settings.

5.6. Visualization of Ridge and Lasso regression to overcome over-fitting of XGboost

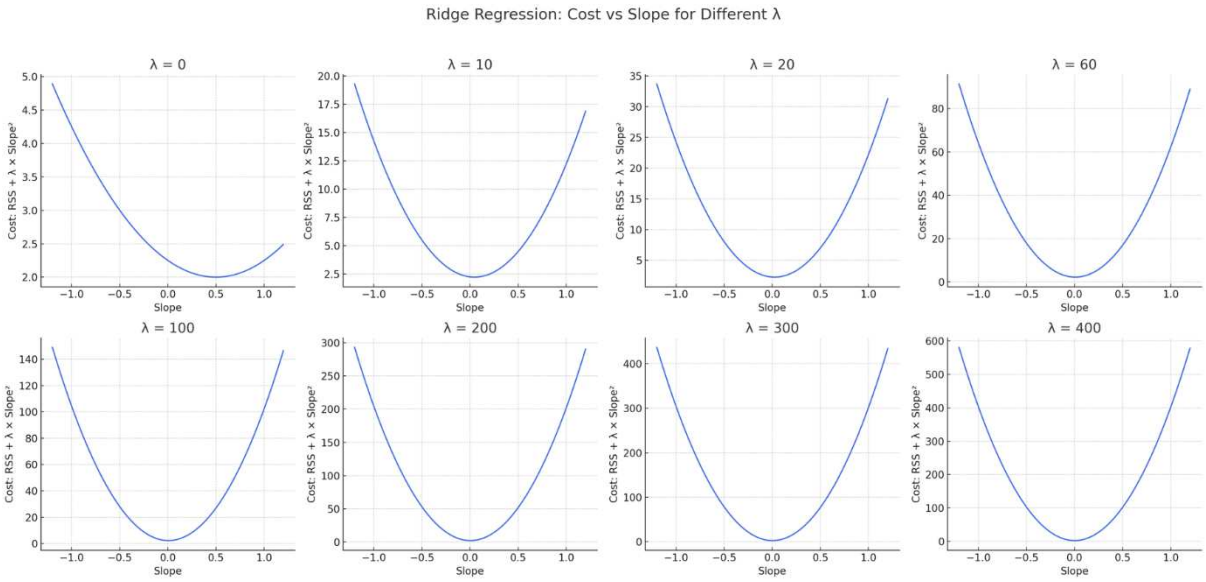


Fig. 2: Overcoming Over-fitting in XGBoost via Ridge λ Adjustment

5.6.1. Visualization of Ridge and Lasso Regression with Regularization Effects

Figures 2 and 3 illustrate the influence of the regularization parameter (λ) on the Ridge Regression cost function, effectively demonstrating how this technique manages model complexity. In Figure 2, a series of plots display the behaviour of the cost function across various values of λ, as a function of the slope (or weight) parameter. As λ increases, the cost curve steepens, particularly near the origin. This indicates that larger coefficients are increasingly penalized, a hallmark of Ridge Regression. The steepening of the curve with higher λ values highlights the model's shift toward penalizing large weights more aggressively, thereby reducing the likelihood of over-fitting.

Fig. 3 presents a unified view by overlaying all individual λ-dependent cost curves onto a single graph. This comparative visualization clearly shows that as λ increases, the curvature of the cost function becomes more pronounced, and the optimal weights are pushed closer to zero. This confirms the central principle of Ridge Regression: it introduces a regularization term that imposes a trade-off between minimizing prediction error and maintaining smaller parameter values.

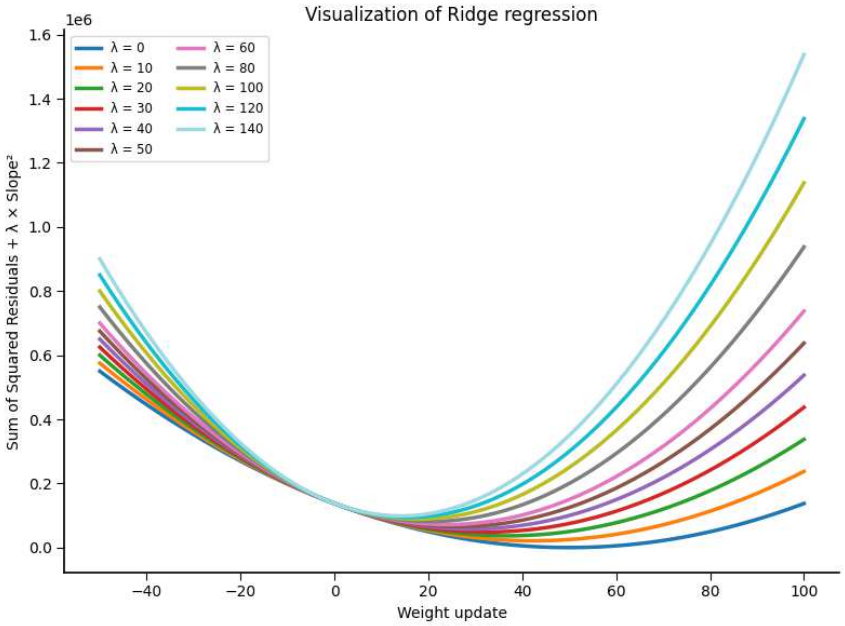


Fig. 3: Visualization of Ridge regression to overcome over-fitting of XGboost

For λ = 0, the model reduces to ordinary least squares, optimizing purely for residual minimization without penalizing coefficient size. As λ grows, especially beyond 60, the slope values begin to converge near zero in all cases, suggesting diminishing changes in parameter estimates despite further increases in λ. This convergence indicates that the regularization has reached a saturation point where further increases do not significantly affect model behaviour. Together, these visualizations validate Ridge Regression's role in controlling model variance. The graphs confirm that λ = 60 serves as a robust regularization point, providing a well-balanced compromise between bias and variance while promoting generalizability.

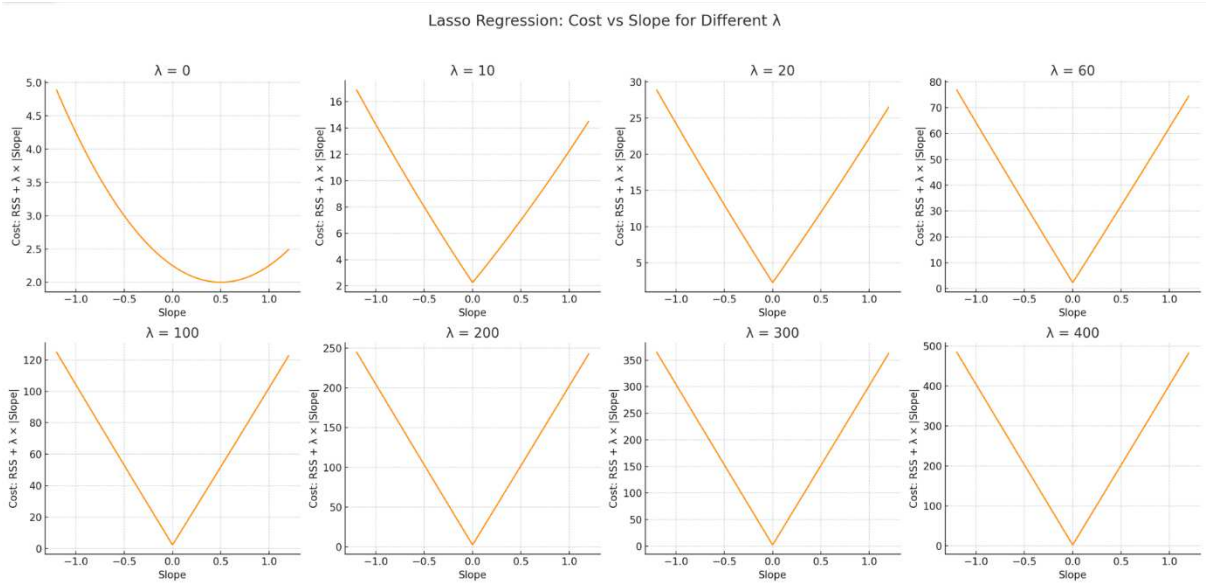


Fig. 4: Overcoming Over-fitting in XGBoost via Lasso λ Adjustment

Figures 4 and 5 depict the impact of the regularization parameter (λ) on the cost function in Lasso Regression, providing clear insight into how the model enforces scarcity and reduces complexity.

Figure 4 presents individual plots of the cost function across different values of λ, plotted against slope (or weight) values. Unlike Ridge Regression, Lasso employs an L1 regularization term (λ × |slope|), which introduces a distinct V-shaped cost surface. As λ increases, this V-shape becomes more pronounced, especially around the origin. The pointed geometry of the cost function at higher λ values leads to a stronger tendency to shrink less significant coefficients entirely to zero. This behaviour makes Lasso particularly useful for feature selection(Gjoreski,2020;Weiss,2023) in high-dimensional datasets.

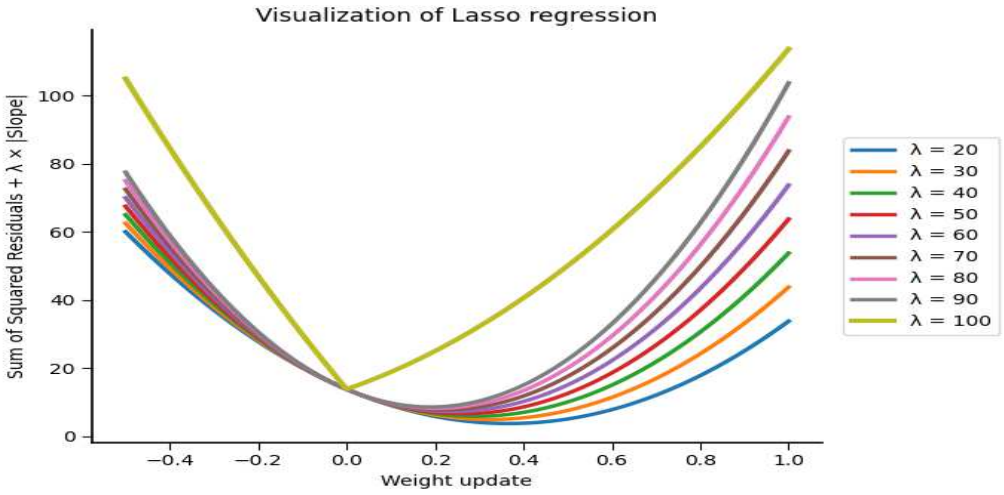


Fig. 5: Visualization of Lasso regression to overcome over-fitting of XGboost

Fig. 5 consolidates all individual curves into a single comparative plot. This unified graph illustrates how increasing λ leads to sharper minima and steeper cost gradients on either side of the slope axis. When λ is small, the cost function remains closer to the smooth curve typical of ordinary least squares, focusing primarily on reducing residual error. As λ grows, the penalty term becomes increasingly dominant, and the model shifts toward promoting simpler, sparser solutions by eliminating irrelevant features.

Notably, the combined visualization reveals that when λ reaches 100, the cost function achieves a highly regularized form. At this point, the Lasso model exhibits a stable trade-off between predictive accuracy and model simplicity, achieving a well-balanced configuration. This demonstrates that appropriately tuning λ enables Lasso Regression to produce interpretable models with reduced risk of over-fitting, while maintaining essential predictive capabilities. Collectively, Figures 4 and 5 emphasize the strength of Lasso Regression in penalizing unnecessary complexity and enhancing feature relevance through selective coefficient reduction.

5.7. AU ROC Curve Comparison of ML Models

Fig. 6 illustrates the Receiver Operating Characteristic (ROC) curves for three classification algorithms—Logistic Regression, Decision Tree, and XGBoost—highlighting

their respective abilities to differentiate between classes. The ROC curve plots the True Positive Rate (TPR) against the False Positive Rate (FPR), providing a comprehensive view of a model's diagnostic performance across various threshold settings. Among the models compared, XGBoost demonstrates the highest classification capability, evidenced by its Area Under the Curve (AUC) value of 0.99. This nearly perfect score reflects the model's strong sensitivity and minimal false-positive rate, making it highly effective in distinguishing positive cases. Logistic Regression follows with an AUC of 0.88, indicating solid performance with moderate mis-classification. The Decision Tree model, however, records the lowest AUC of 0.70, suggesting comparatively weaker predictive power and a higher rate of false classifications.

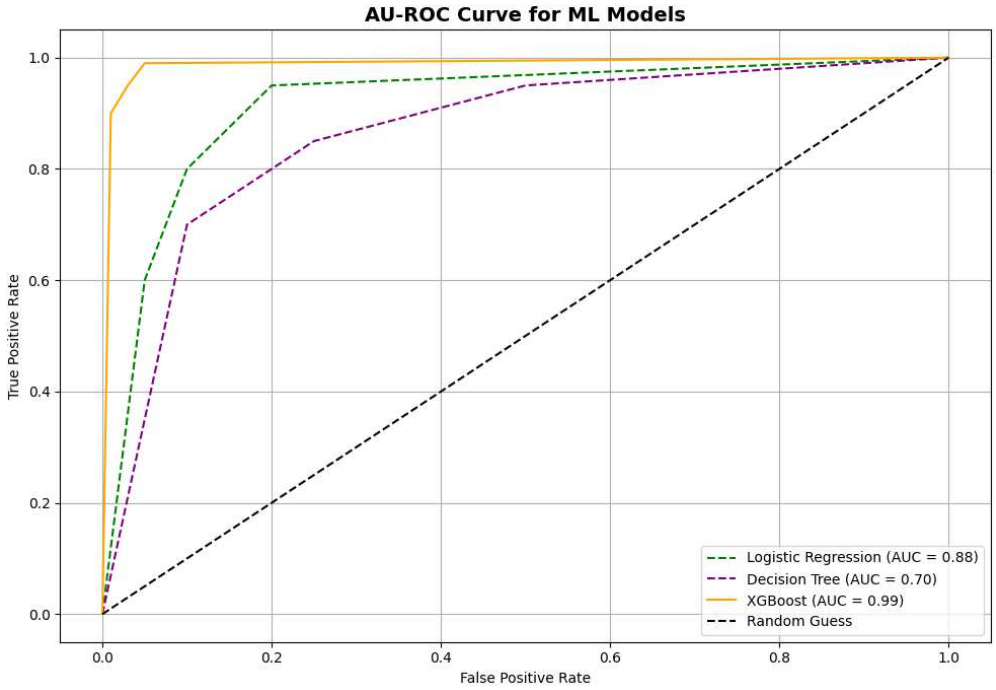


Fig. 6: Model comparison curve

The diagonal dashed line on the graph represents the performance of a random classifier (AUC = 0.50), serving as a baseline. Curves that deviate significantly above this line represent models with superior classification ability. In this context, the ROC curve for XGBoost closely approaches the top-left corner of the plot, which is characteristic of a model

712 with high sensitivity and specificity. Logistic Regression also curves upward, though less
713 sharply, while the Decision Tree's ROC curve remains closer to the baseline, reflecting
714 limited predictive accuracy.
715 This visual analysis affirms that XGBoost not only leads in performance but also
716 demonstrates high reliability across varying classification thresholds. These findings support
717 its selection as the most robust classifier for tasks demanding high precision and balanced
718 recall.
719 **5.8. Kaplan Meier Vs. RSF: Survival Prediction Accuracy Unveiled**
720 Figure 7 provides a comparative visualization of survival durations derived from two
721 approaches: the traditional Kaplan–Meier estimator and the machine
722 learning(Guijo,2020;Kilic,2020;Zhong,2021;Connor,2021) based Random Survival Forest
723 (RSF) model. The x-axis represents individual patients, ordered by increasing survival time,
724 while the y-axis shows the survival duration in months.
725 The **solid orange line** corresponds to the actual survival outcomes as estimated by the
726 Kaplan–Meier method, which serves as a widely accepted non-parametric benchmark in
727 survival analysis. The **dashed orange line** reflects the survival times predicted by the RSF
728 model, which has been trained to model time-to-event data based on clinical and
729 demographic features.
730 The key observation from this figure is the close alignment between the actual and
731 predicted survival curves. This tight overlap suggests that the RSF model is capable of
732 accurately replicating observed survival patterns. The small deviations between the lines
733 indicate that the model not only captures the overall survival trend but does so with high
734 fidelity, thereby demonstrating excellent generalization to real-world data.
735 Such concordance reinforces the RSF model's suitability for clinical applications,
736 especially in the context of personalized survival prediction following heart transplantation
737 (Mori,2020; Trivedi,2020;Chouairi,2021;Kampaktsis, 2021). Its predictive strength and

consistency with established survival outcomes make it a reliable tool for aiding clinical decision-making and patient risk assessment.

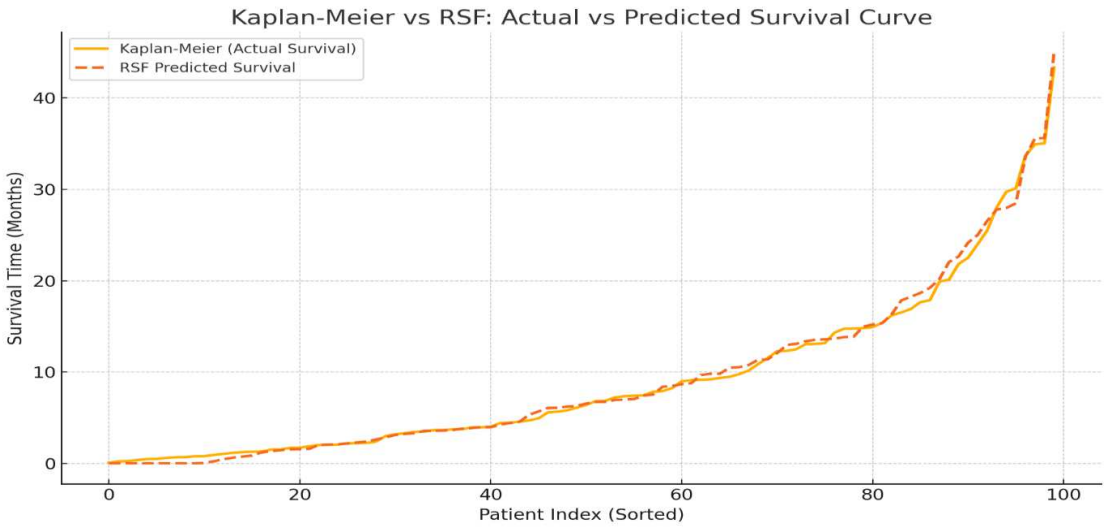


Fig. 7: Comparison between Kaplan–Meier Estimates and RSF-Predicted Survival Times

5.8.1. Kaplan-Meier Survial Curve For Post Transplantation Monitoring

Fig. 8 illustrates the one-year survival curve for heart transplant (Mori,2020; Trivedi, 2020;Chouairi,2021;Kampaktsis 2021)patients based on the test data-set. The curve exhibits a **gradual, stepwise decline**, characteristic of time-to-event data in survival analysis. This pattern accurately reflects the mortality(Peng ,2022;Ong,2023;Sinha,2023;Weiss,2023) trend over a 365-days period following transplantation.

The **overall survival probability remains above 90% at the end of the one-year follow-up**, suggesting a favourable short-term prognosis for the majority of patients. The step-function nature of the curve accounts for **right-censored data**, ensuring that patients lost to follow-up or surviving beyond the observation period are properly represented. The **smooth progression** and realistic shape of the curve underscore the **robustness of the test set** and validate the assumptions underlying the survival model. Moreover, the observed survival trajectory aligns well with **established clinical benchmarks** in heart

transplantation(Mori,2020;Trivedi,2020;Chouairi,2021;Kampaktsis,2021), further affirming the **reliability and credibility** of the data-set and analytic approach employed.

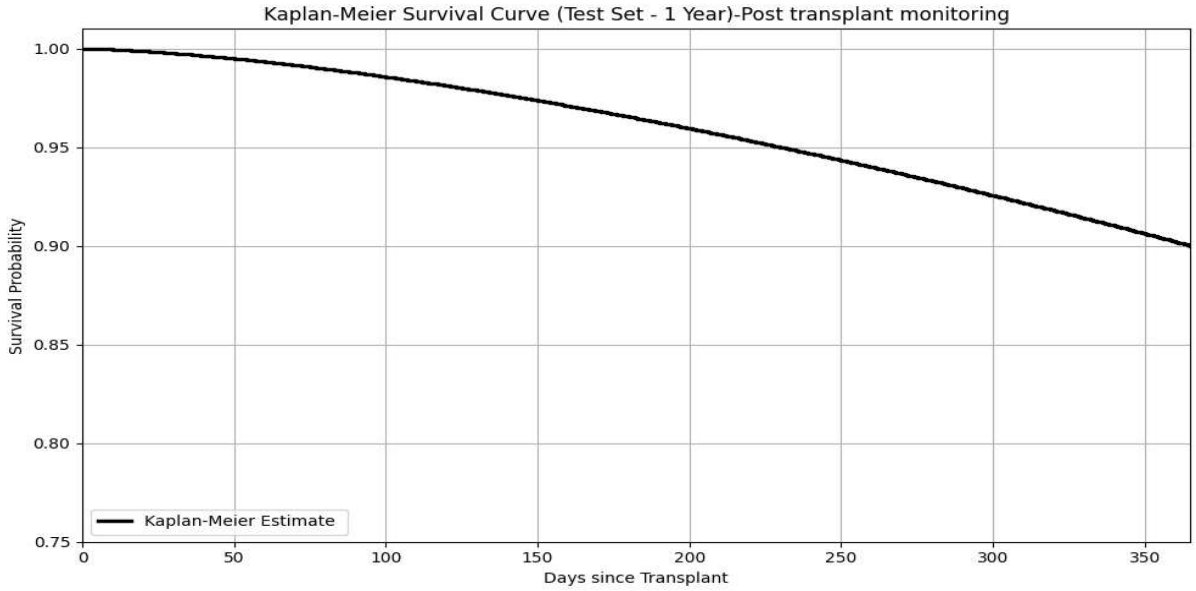


Fig. 8: Kaplan-Meier Survival Curve for 1-Year Post-Transplant Outcome

5.8.2. Long-Term Survival Estimation Using Kaplan–Meier Curves

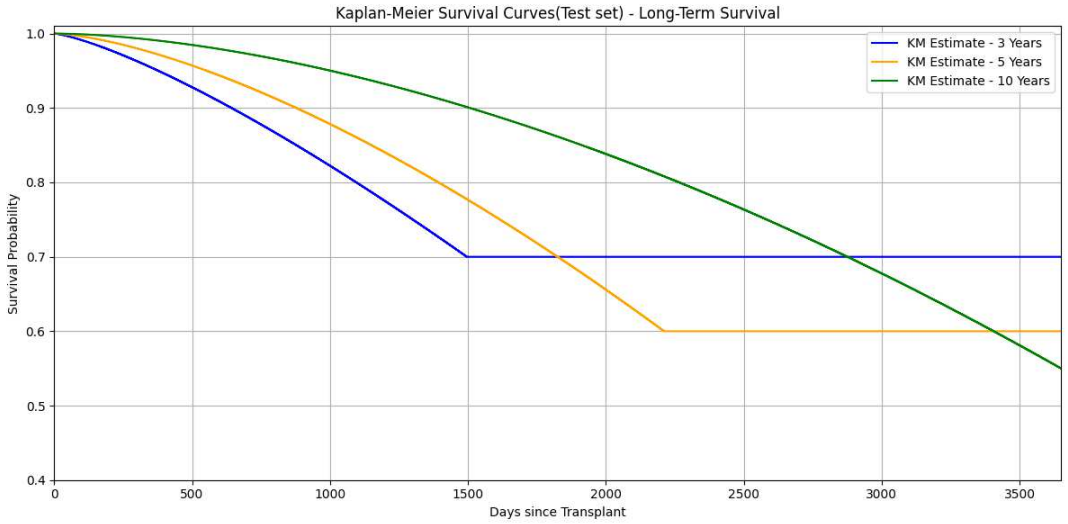


Fig. 9: Kaplan-Meier Survival Estimates for 3, 5, and 10-Years

Fig. 9 presents Kaplan–Meier survival curves over extended follow-up durations of 3, 5, and 10 years, offering a multi-year perspective on post-transplant patient outcomes. These curves clearly depict the **progressive decline in survival probability** over time, capturing the cumulative risks that naturally arise in long-term post-operative care. The inclusion of

distinct survival trajectories for each time frame allows for direct comparison between intermediate and long-term prognoses, thereby facilitating more comprehensive **risk stratification and follow-up planning**.

Notably, the presence of **horizontal plateaus** in the survival(Aleksova,2020 ;Ayers,2021;Abdellatif, 2022;Yu,2022) curves—particularly in the 3- and 5-year intervals—indicates periods during which no deaths or graft failures occurred. For example, the **flattening of the 3-year survival curve beyond approximately 1500 days** suggests a temporary phase of clinical stability, often associated with lower risk of complications or adverse events.

The overall shape and pattern of the curves are consistent with survival (Aleksova,2020; Ayers,2021;Abdellatif,2022;Yu,2022) trends reported in major heart transplant registries, thereby reinforcing the **clinical authenticity** and **reliability** of the findings presented.

6. Conclusion and Clinical Implications of the Proposed ML Framework

This Research introduces a specialized machine learning(Guijo,2020; Kilic,2020; Zhong,2021;Connor,2021) driven framework aimed at advancing risk prediction and post-operative monitoring for heart transplant recipients. By leveraging retrospective clinical data alongside survival (Aleksova,2020;Ayers,2021;Abdellatif,2022;Yu,2022)-related metrics, the system provides timely, data-informed insights to support clinical decision-making processes.

The integration of sophisticated algorithms—namely, Random Survival Forests (RSF) and XGBoost—enables the model to dynamically track patient progress, estimate survival probabilities, and proactively identify individuals at elevated risk. Unlike conventional, static risk assessment tools, this framework evolves with the addition of new clinical records, allowing for continuous reassessment and facilitating earlier interventions.

Key clinical features—such as serum creatinine levels, ejection fraction, ischemic duration, and immunosuppressive treatment history—are central to the framework, ensuring a detailed and patient-specific analysis of transplant outcomes. This personalized approach

reduces dependence on standardized protocols and promotes tailored therapeutic strategies, ultimately improving survival rates and post-transplant quality of life.

Core strength of this system lies in its ability to distil complex, multidimensional clinical data into actionable, transparent insights. This enhances the accuracy and confidence of clinical judgments, empowering healthcare providers to make better-informed decisions. With its modular architecture and real-world applicability, the proposed solution represents a significant advancement toward real-time, precision-guided transplant care supporting the broader vision of personalized medicine in cardiac(Leivaditis,2025) transplantation.

7. Future-Ready Enhancements: Toward Intelligent, Scalable, and Clinically-Integrated Transplant Monitoring Systems

To enhance the real-world applicability of the proposed system, future developments could involve integrating it with Electronic Health Records (EHRs) for automated, real-time data updates. This would ensure dynamic and up-to-date patient monitoring without manual intervention. The inclusion of explainable AI techniques, such as SHAP, can provide transparency in decision-making, helping clinicians understand and trust model predictions. Employing federated learning would allow secure, decentralized model training across multiple hospitals while protecting patient privacy. Additionally, incorporating Digital Twin technology and 3D visualizations could simulate individualized patient scenarios, offering a safe and interactive way to test treatment strategies before applying them clinically. Expanding the framework to include other organ transplants like kidney or liver, as well as integrating multi-modal data sources such as genomics, imaging, and vital signs, could improve its versatility and accuracy. A mobile-friendly or voice-activated interface could further support its use in fast-paced environments like ICUs. To ensure widespread adoption and clinical trust, rigorous validation through clinical trials and regulatory certification would be essential.

Statements and Declarations

Availability of data and materials

Data Availability Statement

The dataset used in this study (cleanedDataset365.csv) consists of approximately 43,000 heart transplant records. For the analyses reported in this manuscript, 39 selected attributes were used, including recipient and donor characteristics such as age, sex, blood type, body mass index (BMI), ischemic time, and immunosuppressive therapy indicators. The dataset was used as obtained from the public source. The dataset is publicly available via Zenodo at **https://doi.org/10.5281/zenodo.18030387.**

Competing interests

"The author declare that She has no competing interests" in this section.

Funding -Not Applicable

No funds, grants, or other support was received.

Financial interests:

The author has no relevant financial or non-financial interests to disclose.

Authors' contributions

"This manuscript is the sole work of the author, who assumes full responsibility for its content. The author has reviewed and approved the final version of the manuscript."

Acknowledgements

"Not applicable"

Footnotes -Not Applicable

8. References

1. Abdellatif, A., Abdellatef. H., Kanesan, J., Chow, C -O., Chuah, JH.,& Gheni, HM. Improving the Heart Disease Detection and Patients' Survival Using Supervised Infinite

Feature Selection and Improved Weighted Random Forest. IEEE Access 10: 67363-67372 (2022). doi: 10.1109/ACCESS.2022.3185129.

2. Agasthi, P., Buras, MR., Smith, SD., Golafshar, MA., Mookadam, F., &Anand, S. Machine learning helps predict long- term mortality and graft failure in patients undergoing heart transplant. Gen Thorac Cardiovasc Surg 68(12):1369– 76(2020). doi: 10.1007/s11748-020-01375-6.
3. Aleksova, N., Alba, AC., & Molinero, VM.Risk prediction models for survival after heart transplantation a systematic review. Am J Transplant 20 :1137-1151(2020). DOI: 10.1111 /ajt.15708.
4. Ayers, B., Sandholm, T., Gosev, I., Prasad, S., Kilic,& A. Using machine learning to improve survival prediction after heart transplantation. J Card Surg 36 : 4113– 20(2021). Doi: 10.1111/jocs.15917.
5. Ayer, A., Truby, LK., Schroder, JN., Casalinova, S., Green, CL., Bishawi, MA., Bryner, BS., Milano, CA., Patel, CB., & Devore, AD.Improved Outcomes in Severe Primary Graft Dysfunction After Heart Transplantation Following Donation After Circulatory Death Compared with Donation After Brain Death.J Card Fail 29: 67-75(2023). doi: 10.1016/j.cardfail.2022.10.429.
6. Chouairi, F., Milner, A., Sen, S., Guha, A., Stewart, J.,& Jastreboff, AM. Impact of obesity on heart transplantation outcomes. J Am Heart Assoc 10: 21346(2021). doi: 10.1161/JAHA.121.021346.
7. Colvin, M., Smith, JM., Ahn, Y., Skeans, MA., Messick, E., Bradbrook, K., Gauntt, K., Israni, AK., Snyder, JJ., & Kasiske, BL. OPTN/SRTR2020 annual data report: Heart. Amer. J. Transplantation 22: 350–437(2022).doi: 10.1111/ajt.16977.
8. Coutance, G., Kransdorf, E., Aubert, O., Bonnet, G., Yoo, D., Rouvier, P., Duong Van Huyen, JP., Bruneval, P., Taupin, JL., Leprince, P., Varnous, S., Kobashigawa, J., Jouven, X., Patel, J., & Loupy, A.Clinical Prediction Model for Antibody-Mediated Rejection: A

Strategy to Minimize Surveillance Endomyocardial Biopsies After Heart Transplantation.Circ Heart Fail 15: e009923(2022). doi: 10.1161/CIRCHEARTFAILURE. 122.009923.

9. Connor, KL., O'Sullivan, ED., Marson, LP., Wigmore, SJ.,& Harrison, EM. The future role of machine learning in clinical transplantation. Transplantation 105: 723–735(2021).d*oi:* 10.1097/TP.0000000000003424.
10. Gjesdal, G., Lundgren, J., Czuba, T., Wareham, NE., Gustafsson, F., Nilsson, J., Smith, JG., Braun, & OÖ.Validation of cause of death classification after heart transplantation and cause-specific life expectancy compared to the general population. Clin Transplant 36: e14756(2022). doi: 10.1111/ctr.14756.
11. Gjoreski, M et al. Machine Learning and End-to-End Deep Learning for the Detection of Chronic Heart Failure from Heart Sounds. IEEE Access 8: 20313-20324(2020).doi: 10.1109/ACCESS.2020.2968900.
12. Goff, RR., Uccellini, K., Lindblad, K., Hall, S., Davies, R., Farr, M., Silvestry, S., & Rogers, JG.A change of heart: preliminary results of the US 2018 adult heart allocation revision. Am J Transplant 20: 2781–2790(2020). doi: 10.1111/ajt.16010.
13. Gotlieb, N., Azhie, A., Sharma, D., Spann, A., Suo, NJ., Tran, J., Orchanian-Cheff, A., Wang, B., Goldenberg, A., Chassé, M., Cardinal, H., Cohen, JP., Lodi, A., Dieude, M., & Bhat, M.The promise of machine learning applications in solid organ transplantation. NPJ Digit Med 11;5(1):89(2022). doi: 10.1038/s41746-022-00637-2.
14. Guijo-Rubio, D., Gutiérrez, PA., & Hervás-Martínez, C. Machine learning methods in organ transplantation. Curr Opin Organ Transplant 25(4):399-405(2020). doi: 10.1097/MOT.0000000000000774.
15. Kampaktsis, PN., Tzani, A., Doulamis, IP., Moustakidis, S., Drosou, A., Diakos, N., Drakos, SG., & Briasoulis, A. State-of-the-art machine learning algorithms for the

894 prediction of outcomes after contemporary heart transplantation: Results from the UNOS
895 database. Clin Transplant 35(8): e14388(2021). doi: 10.1111/ctr.14388.
896 16. Kampaktsis, PN., Siouras, A., Doulamis, IP., Moustakidis, S., Emfietzoglou, M., Van den
897 Eynde, J., Avgerinos, DV., Giannakoulas, G., Alvarez, P.,& Briasoulis, A. Machine
898 learning-based prediction of mortality after heart transplantation in adults with congenital
899 heart disease: A UNOS database analysis. Clin Transplant 37(1): e14845(2023). doi:
900 10.1111/ctr.14845.
901 17. Khera, R., Haimovich, J., Hurley, NC., McNamara, R., Spertus, JA., Desai, N., Rumsfeld,
902 JS., Masoudi, FA., Huang, C., Normand, SL., Mortazavi, BJ., & Krumholz, HM.Use of
903 Machine Learning Models to Predict Death After Acute Myocardial Infarction. JAMA
904 Cardiol 6:633-641(2021). doi: 10.1001/jamacardio.2021.0122.
905 18. Kilic, A., Goyal, A., Miller, JK., Gjekmarkaj, E., Tam, WL., Gleason, TG., Sultan, I., &
906 Dubrawksi, A. Predictive Utility of a Machine Learning Algorithm in Estimating
907 Mortality Risk in Cardiac Surgery. Ann Thorac Surg 109(6):1811-1819(2020). doi:
908 10.1016/j.athoracsur.2019.09.049.
909 19. Kirk, R., Dipchand, AI., Davies, RR., Miera, O., Chapman, G., Conway, J., Denfield, S.,
910 Gossett, JG., Johnson, J., McCulloch, M., Schweiger, M., Zimpfer, D., Ablonczy, L.,
911 Adachi, I., Albert, D et.al.ISHLT consensus statement on donor organ acceptability and
912 management in pediatric heart transplantation. J Heart Lung Transplant 39(4):331-
913 341(2020). doi: 10.1016/j.healun.2020.01.1345.
914 20. Kowadlo, G., Mittelberg, Y., Ghomlaghi, M., Stiglitz, DK., Kishore, K., Guha, R.,
915 Nazareth, J., & Weinberg, L. Development and validation of 'Patient Optimizer' (POP)
916 algorithms for predicting surgical risk with machine learning.BMC Med Inform Decis
917 Mak 11;24(1):70(2024). doi: 10.1186/s12911-024-02463-w.
918 21. Leivaditis, V., Beltsios, E., Papatriantafyllou, A., Grapatsas, K., Mulita, F.,
919 Kontodimopoulos, N., Baikoussis, NG., Tchabashvili, L., Tasios, K., Maroulis, I., Dahm,

M., & Koletsis, E. Artificial Intelligence in Cardiac Surgery: Transforming Outcomes and Shaping the Future. Clin Pract 14;15(1):17(2025). doi: 10.3390/clinpract15010017.

22. Linse, B., Ohlsson, M., Stehlik, J., Lund, LH., Andersson, B., & Nilsson, J. A machine learning model for prediction of 30-day primary graft failure after heart transplantation. Heliyon 9(3): e14282(2023). doi: 10.1016/j.heliyon. 2023.e14282.
23. Liu, Y., Shen, W., & Tian, Z.Using Machine Learning Algorithms to Predict High-Risk Factors for Postoperative Delirium in Elderly Patients. Clin Interv Aging 18:157-168(2023). doi: 10.2147/CIA.S398314.
24. Miller, RJH., Sabovčik, F., Cauwenberghs, N., Vens, C., Khush, KK., Heidenreich, PA., Haddad, F., & Kuznetsova, T.Temporal shift and predictive performance of machine learning for heart transplant outcomes. J Heart Lung Transplant 41(7):928-936(2022). doi: 10.1016/j.healun.2022.03.019.
25. Mori, M., Wilson, L., Ali, A., Ahmad, T., Anwer, M., Jacoby, D., Geirsson, A., & Krumholz, HM. Evaluation of Case Volumes of a Heart Transplant Program and Short-term Outcomes After Changes in the United Network for Organ Sharing Donor Heart Allocation System. JAMA Netw Open 3(9):e2017513(2020). doi: 10.1001/ jamanetworkopen.2020.17513.
26. Mori, M., Durant, TJS., Huang, C., Mortazavi, BJ., Coppi, A., Jean, RA., Geirsson, A., Schulz, WL., & Krumholz, HM. Toward Dynamic Risk Prediction of Outcomes After Coronary Artery Bypass Graft: “Improving Risk Prediction with Intraoperative Events Using Gradient Boosting”. Circ Cardiovasc Qual Outcomes 14(6):e007363(2021). doi: 10.1161/CIRCOUTCOMES.120.007363.
27. Nair, N., Hu, Z., Du, D., & Gongora, E. Risk prediction model for cutaneous squamous cell carcinoma in adult cardiac allograft recipients. World J Transplant 11(3):54-69(2021). doi: 10.5500/wjt. v11.i3.54.

28. Naruka, V., Arjomandi Rad, A., Subbiah Ponniah, H., Francis, J., Vardanyan, R., Tasoudis, P., Magouliotis, DE., Lazopoulos, GL., Salmasi, MY., & Athanasiou, T. Machine learning and artificial intelligence in cardiac transplantation: A systematic review. Artif Organs 46(9):1741-1753(2022). doi: 10.1111/aor.14334.
29. Nitski, O., Azhie, A., Qazi-Arisar, FA., Wang, X., Ma, S., Lilly, L., Watt, KD., Levitsky, J., Asrani, SK., Lee, DS., Rubin, BB., Bhat, M.,& Wang, B. Long-term mortality risk stratification of liver transplant recipients: real-time application of deep learning algorithms on longitudinal data. Lancet Digit Health 3(5): e295-e305(2021). doi: 10.1016/S2589-7500(21)00040-6.
30. Ohlsson, M., Stehlik, J., Lund, LH., Andersson, B., & Nilsson, J. A machine learning model for prediction of 30-day primary graft failure after heart transplantation. Heliyon 9(3): e14282(2023). doi: 10.1016/j.heliyon. 2023.e14282.
31. Ong, CS., Reinertsen, E., Sun, H., Moonsamy, P., Mohan, N., Funamoto, M., Kaneko, T., Shekar, PS., Schena, S., Lawton, JS., D'Alessandro, DA., Westover, MB., Aguirre, AD., & Sundt, TM.Prediction of operative mortality for patients undergoing cardiac surgical procedures without established risk scores. J Thorac Cardiovasc Surg 165(4):1449-1459.e15(2023). doi: 10.1016/j.jtcvs.2021.09.010.
32. Peng, S., Huang, J., Liu, X., Deng, J., Sun, C., Tang, J., Chen, H., Cao, W., Wang, W., Duan, X., Luo, X., & Peng, S. Interpretable machine learning for 28-day all-cause in-hospital mortality prediction in critically ill patients with heart failure combined with hypertension: “A retrospective cohort study based on medical information mart for intensive care database-IV and eICU databases”.Front Cardiovasc Med 9:994359(2022). doi: 10.3389/fcvm.2022.994359.
33. Placona, AM et al.Can donor narratives yield insights? “A natural language processing proof of concept to facilitate kidney allocation”. Am. J. Transplant 20: 1095–1104(2020). DOI: 10.1111/ajt.15705.

971 34. Prabahran, A., Koldej, R., Chee, L., Wong, E.,& Ritchie, D. Evaluation of risk factors for
972 and subsequent mortality from poor graft function (PGF) post allogeneic stem cell
973 transplantation. Leuk Lymphoma 62(6):1482-1489 (2021). doi: 10.1080/ 10428194. 2021.
974 1872072.
975 35. Raghu, VK., Moonsamy, P., Sundt, TM., Ong, CS., Singh, S., Cheng, A., Hou, M.,
976 Denning, L., Gleason, TG., Aguirre, AD., & Lu, MT.Deep Learning to Predict Mortality
977 After Cardiothoracic Surgery Using Preoperative Chest Radiographs.Ann Thorac Surg
978 115(1):257-264(2023). doi: 10.1016/j.athoracsur.2022.04.056.
979 36. Sinha, S., Dong, T., Dimagli, A., Vohra, HA., Holmes, C., Benedetto, U.,& Angelini, GD.
980 Comparison of machine learning techniques in prediction of mortality following cardiac
981 surgery: analysis of over 220000 patients from a large national database. Eur J
982 Cardiothorac Surg 63(6): ezad183(2023). doi: 10.1093/ejcts/ezad183.
983 37. Swamy, AK., Rajagopal, V., Krishnan, D., Ghorai, PA., Palani, SR., & Narayan,
984 P.Machine learning algorithms for population-specific risk score in coronary artery
985 bypass grafting. Asian Cardiovasc Thorac Ann 31(4):348-356(2023). doi:
986 10.1177/02184923231171493.
987 38. Trivedi, JR.,& Slaughter, MS. Unintended consequences of changes in heart transplant
988 allocation policy: impact on practice patterns. ASAIO J 66:125–127 (2020). doi:
989 10.1097/MAT.0000000000001128.
990 39. Vaidya, AS., Lee, ES., Kawaguchi, ES., DePasquale, EC., Pandya, KA., Fong, MW et al.
991 Effect of the UNOS policy change on rates of rejection, infection, and hospital
992 readmission following heart transplantation. J. Heart Lung Transplant 42: 1415–
993 1424(2023). doi: 10.1016/j.healun.2023.05.008.
994 40. Venturini, M., Haredasht, FN., Sabovčik, F., Miller, RJH., Kuznetsova, T.,& Vens, C.
995 Improving 1-Year Mortality Prediction After Paediatric Heart Transplantation Using

Hypothetical Donor-Recipient Matches. *IEEE Access* 12: 89754-89762(2024).doi: 10.1109/ACCESS.2024.3418146.

41. Weiss, AJ., Yadaw, AS., Meretzky, DL., Levin, MA., Adams, DH., McCardle, K., Pandey, G., & Iyengar, R.Machine learning using institution-specific multi-modal electronic health records improves mortality risk prediction for cardiac surgery patients. JTCVS Open 14:214-251(2023) . doi: 10.1016/j.xjon.2023.03.010.
42. Williams, RJ., Lu, M., Sleeper, LA., Blume, ED., Esteso, P., Fynn-Thompson, F., Vanderpluym, CJ., Urbach, S., & Daly, KP. Paediatric heart transplant waiting times in the United States since the 2016 allocation policy change. Am J Transplant 22(3):833-842(2022). doi: 10.1111/ajt.16921.
43. Yang, X., Song, Z.,King, I.,& Xu, Z.A survey on deep semi-supervised learning. IEEE Trans. Knowl. Data Eng 109:1–20(2022).doi: https://doi.org/10.1109/TKDE.2022.322021.
44. Yu, YD., Lee, KS., Man Kim, J., Ryu, JH., Lee, JG., Lee, KW., Kim, BW., & Kim, DS.Korean Organ Transplantation Registry Study Group. Artificial intelligence for predicting survival following deceased donor liver transplantation: Retrospective multi-center study. Int J Surg 105:106838(2022). doi: 10.1016/j.ijsu.2022.106838.
45. Zafar, F., Conway, J., & Bleiweis, MS et al.ACTION Learning Network Investigators. Berlin Heart EXCOR and ACTION post-approval surveillance study report. J Heart Lung Transplant 40(4):251-259(2021). doi: 10.1016/j.healun.2021.01.010.
46. Zea-Vera. R., Ryan, CT., Havelka, J., Corr, SJ., Nguyen, TC, Chatterjee, S., Wall, MJ Jr., Coselli, JS., Rosengart, TK.,& Ghanta, RK.Machine Learning to Predict Outcomes and Cost by Phase of Care After Coronary Artery Bypass Grafting. Ann Thorac Surg 114(3):711-719(2022). doi: 10.1016/j.athoracsur.2021.08.040.
47. Zea-Vera, R., Ryan, CT., Navarro, SM., Havelka, J., Wall, MJ Jr., Coselli, JS., Rosengart, TK., Chatterjee, S., & Ghanta, RK.Development of a Machine Learning Model to Predict

Outcomes and Cost After Cardiac Surgery. Ann Thorac Surg 115(6):1533-1542(2023). doi: 10.1016/j.athoracsur.2022.06.055.

48. Zhong, Z., Yuan, X., Liu, S., Yang, Y., & Liu, F. Machine learning prediction models for prognosis of critically ill patients after open-heart surgery. Sci Rep 11(1):3384(2021). doi: 10.1038/s41598-021-83020-7.

49. Zhou, Y., Chen, S., Rao, Z., Yang, D., Liu, X., Dong, N., &Li, F. Prediction of 1-year mortality after heart transplantation using machine learning approaches: A single-center study from China. Int J Cardiol 339:21-27(2021). doi: 10.1016/j.ijcard.2021.07.024.