

RESEARCH

Open Access



# Developing clinical prognostic models to predict graft survival after renal transplantation: comparison of statistical and machine learning models

Getahun Mulugeta<sup>1\*</sup>, Temesgen Zewotir<sup>2</sup>, Awoke Seyoum Tegegne<sup>1</sup>, Mahteme Bekele Muleta<sup>3</sup> and Leja Hamza Juhar<sup>3</sup>

## Abstract

**Introduction** Renal transplantation is a critical treatment for end-stage renal disease, but graft failure remains a significant concern. Accurate prediction of graft survival is crucial to identify high-risk patients. This study aimed to develop prognostic models for predicting renal graft survival and compare the performance of statistical and machine learning models.

**Methodology** The study utilized data from 278 renal transplant recipients at the Ethiopian National Kidney Transplantation Center between September 2015 and February 2022. To address the class imbalance of the data, SMOTE resampling was applied. Various models were evaluated, including Standard and penalized Cox models, Random Survival Forest, and Stochastic Gradient Boosting. Prognostic predictors were selected based on statistical significance and variable importance.

**Results** The median graft survival time was 33 months, and the mean hazard of graft failure was 0.0755. The 3-month, 1-year, and 3-year graft survival rates were found to be 0.979, 0.953, and 0.911, respectively. The Stochastic Gradient Boosting (SGB) model demonstrated the best discrimination and calibration performance, with a C-index of 0.943 and a Brier score of 0.000351. The Ridge-based Cox model closely followed the SGB model's prediction performance with better interpretability. The key prognostic predictors of graft survival included an episode of acute and chronic rejections, post-transplant urological complications, post-transplant nonadherence, blood urea nitrogen level, post-transplant regular exercise, and marital status.

**Conclusions** The Stochastic Gradient Boosting model demonstrated the highest predictive performance, while the Ridge-Cox model offered better interpretability with a comparable performance. Clinicians should consider the trade-off between prediction accuracy and interpretability when selecting a model. Incorporating these findings into the clinical practice can improve risk stratification and personalized management strategies for kidney transplant recipients.

\*Correspondence:  
Getahun Mulugeta  
gech.marr@gmail.com

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

**Keywords** Renal transplant, Graft survival, SMOT oversampling, Prognostic models, Statistical models, Machine learning models

## Introduction

Renal transplantation is a life-saving treatment that provides improved quality of life and long-term survival for those suffering from end-stage renal disease (ESRD). However, despite its benefits, graft failure is still a significant concern and a major contributor to morbidity and mortality in transplant recipients. Graft failure can occur due to various factors, such as acute rejection, chronic rejection, infection, and other complications. When graft failure occurs, patients may need to resume dialysis or undergo re-transplantation, both of which are associated with poorer outcomes and increased healthcare costs [1].

In the era of personalized and precision medicine, predicting the prognosis of diseases has become vital for patient management by healthcare personnel [2]. Accurate prediction of renal graft survival is critical since it can aid in identifying individuals at high risk of graft failure for closer monitoring and more aggressive therapy. This can potentially improve transplant outcomes and reduce healthcare costs by preventing or delaying graft failure [3]. Graft survival prediction after renal transplantation is considered one of the most challenging and vital aspects of modern medicine [4]. It is difficult as it depends on various factors associated with the donor, transplant, and recipient, and their importance changes over time and per outcome measure [5].

Survival prognostic models estimate the likelihood of an event in a specified timeframe. There are two common approaches used to estimate survival probability: machine learning and statistical modeling [6]. Machine learning models, a subset of statistical models, use algorithms to learn from data and make predictions or decisions without explicit programming. Statistical models, on the other hand, are based on statistical theory and use mathematical equations to model the relationship between variables [7]. Machine learning models are often used when the relationship between variables is complex and not well understood, while statistical models are used when the relationship between variables is well understood and can be modeled using mathematical Eq. [8].

Statistical models, such as Cox models, are widely used and well-understood in survival analysis. They are relatively straightforward to implement and interpret and can provide estimates of the effects of individual variables on the outcome [9]. However, statistical models may not capture complex interactions between variables and may not handle high-dimensional data or non-linear relationships. Machine learning models, instead, have the potential to capture complex interactions and provide more

accurate and personalized predictions by considering the unique characteristics of each study subject. They can handle high-dimensional data and non-linear relationships but may be more difficult to interpret and require more computational resources [6].

To date, numerous studies have employed statistical methods, particularly Cox regression, to develop prognostic models that predict renal graft survival. Although some studies have utilized machine learning methods, some of these models may be overly complex and prone to overfitting. As a result, when tested on new and independent datasets, these models may exhibit poor generalizability and performance [10]. However, it is still controversial which methods among machine learning algorithms and conventional statistical modeling can achieve better performance in survival analysis, particularly in the field of transplant medicine [11]. This study addresses the shortcomings of previous research, which primarily compared the standard Cox model with random survival forests while overlooking essential sociodemographic factors, such as regular physical exercise and marital status, that significantly impact graft survival in the Ethiopian context. In contrast, the current study provides a thorough comparison of models and predictors for graft survival. We effectively tackle overfitting issues and improve prediction performance using techniques like cross-validation, pre-feature selection, hyperparameter tuning, and the use of penalized and tree-based ensemble models.

Accordingly, the current study aimed to compare the performance of statistical and machine learning models for predicting renal graft survival among renal transplant recipients in the Ethiopian National Kidney Transplantation Center (ENKTC). The ENKTC is the sole renal transplant center in Ethiopia and as a newly established facility, it has not yet conducted extensive studies on transplant complications, including graft failure.

## Methods and participants

### Source of data and study design

The institutional retrospective study was conducted on 278 kidney transplant recipients at the Ethiopian National Kidney Transplantation Center between September 2015 and February 2022. The current study included transplant recipients who had three or more follow-up visits in the defined period. The data, including epidemiological, laboratory, and clinical histories, was extracted from patient follow-up charts and medical records. There were no missing values in the data, as the records were crosschecked and any gaps were filled by

contacting patients. The data extraction tool was based on renal post-transplant follow-up guidelines to ensure all variables were accounted for.

#### **Methodological strategies to fix data limitations**

Given that our transplant center is new, established in September 2015, we are working with a relatively small cohort of transplant recipients, comprising only 278 cases, and we have observed a limited number of graft failures, a total of 21 cases. We recognize that the small sample size and the imbalanced nature of our data may impact the reliability of our findings. To enhance both the validity and reliability of our results, we have implemented several methodological strategies:

#### **Pre-feature selection**

We prioritize identifying and incorporating only the most relevant features in our analysis to enhance model effectiveness and directly address our research questions. Using statistical methods and expert knowledge, we conducted pre-feature selection using the `uni.selection` function in R, which performs univariate Cox regression for each predictor in the training dataset. A variable is deemed significant if its p-value is less than or equal to 0.05. After refining for near-zero variance and correlated predictors, we selected 39 out of 54 features. Among these, the univariate Cox and domain experts confirmed the top 10 candidate prognostic features. The final features for developing clinical prognostic models include episodes of chronic and acute rejection, urological complications, nonadherence, glomerulonephritis, post-transplant admissions, blood urea nitrogen levels, delayed graft function, regular physical exercise, marital status [12].

#### **Cross-validation**

We use cross-validation techniques to evaluate model performance across different subsets of our data. This method increases the robustness of our findings by minimizing dependence on any single data partition and enhancing overall reliability [13]. Therefore, we consistently performed 5-fold cross-validation for all model development.

#### **Hyperparameter tuning**

To improve model accuracy, we conduct hyperparameter tuning, which involves optimizing the parameters that dictate the model's behavior. This process plays a crucial role in enhancing predictive performance, particularly in datasets with limited events [14]. Consequently, we tuned the hyperparameters for penalized Cox models, random survival forests, and stochastic gradient boosting models using the random search method, which is presented in the study's results.

#### **Penalized models**

Utilizing penalized models allows us to mitigate potential overfitting issues arising from the limited number of events. These models introduce penalties for complexity, promoting simpler models that generalize better to unseen data [15]. Accordingly, the penalized versions of the Cox model were considered.

#### **SMOTE resampling**

To address a class imbalance in our dataset, where graft failures represent a minority class, we apply SMOTE (Synthetic Minority Over-sampling Technique). This technique enhances the model's learning capabilities by generating synthetic samples from the minority class instead of merely reproducing the existing minority class samples [16]. After applying SMOTE, the class distribution in the training dataset was significantly more balanced. The majority-to-minority class ratio in the training dataset reduced from 12:1 to 1.7:1. This reduction in class imbalance through SMOTE aimed to improve the model's ability to learn the decision boundaries between the two classes, leading to better generalization performance on the hold-out testing data set.

#### **Tree-based ensemble models**

We leverage tree-based ensemble models that are known for their resilience against overfitting and their ability to capture complex interactions among features. These models effectively address the challenges posed by imbalanced data and small sample sizes. Their robustness against overfitting, ability to generalize well, and adaptability through parameter tuning make them a preferred choice in various machine learning tasks. The effectiveness of tree-based ensembles can be further enhanced by combining them with resampling techniques like the SMOTE method. This combination can help balance the dataset while leveraging tree-based models' strengths [17].

Despite the constraints of our data, we aim to derive significant insights from it by employing these methodological tools. This rigorous approach justifies the adequacy of our sample size and deepens our understanding of graft failure dynamics within the context of our emerging transplant center. Moreover, the study adhered to the TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) guidelines to ensure transparent reporting of our prediction model [18].

#### **Experimental setups**

The experiment starts by cleaning, preprocessing, and exploration of the dataset. In this study, we utilized baseline records for longitudinal predictors. Continuous predictors were standardized using the mean-centering

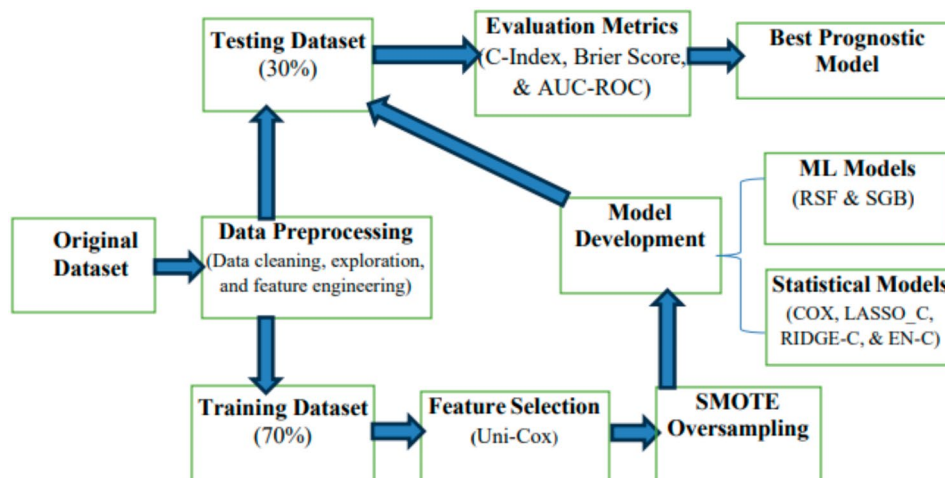
method, while categorical predictors were appropriately coded. Predictors with high correlation and near-zero variance were excluded from the dataset. The preprocessed dataset was divided into a training (70%) and validation set (30%), while maintaining the event/censoring proportion of the original data. Before the model training phase, we conducted data-driven pre-feature selection and applied class imbalance handling techniques to mitigate the risk of overfitting. The refined predictors and balanced training data were used throughout the model development. Models were trained using 5-fold cross-validation and hyperparameter tuning. The predictive performance of the candidate models was evaluated using the hold-out testing dataset, which remained unseen and unmodified during the training and resampling procedures. The appropriate prediction performance evaluation metrics, such as concordance index, Brier score, and area under the ROC curve were used to evaluate and compare the effectiveness of the prognostic models to predict renal graft survival. All experiments were executed on the R platform and R packages like; “survival”, “RandomForestSRS”, “glmnet”, and “gbm” were used to train and validate the specified statistical and machine learning models. The process of selecting significant predictors for graft survival was conducted by assessing the statistical significance and relative importance of candidate variables. Variables that were selected as significant or important by most models were considered potential predictors of graft survival. Ultimately, the best-performing model was reported as a calibrated clinical prognostic model to predict graft survival among renal transplant recipients. Figure 1 illustrates the workflow of the study’s computations.

### Clinical endpoint (outcome)

The clinical outcome for this study is renal graft survival where graft failure is a primary event of interest. Patients who died with a functioning graft were not considered positive for graft failure but were included in all models and censored at their death time. The time to graft failure was measured in months from the date of transplantation. Given the retrospective nature of the study, we defined patients who did not experience graft failure during the study period (September 2015 to February 2022) as right-censored. Moreover, patients who died with a functioning graft during the study period were also treated as right-censored observations. The time of censorship (death or end of the study period) was still included in the analysis. This retrospective approach preserves valuable survival information and enhances the integrity of our survival analysis, ultimately improving the model’s predictive accuracy.

### Consideration of survival prognostic models

Based on the nature of the data and the motive of the current study various statistical and machine-learning models assumed to be effective in this scenario were selected. From the statistical perspective, standard and regularized Cox models such as Lasso-based Cox, Ridge-based Cox, and Elastic net-based Cox were chosen based on their interpretability and appropriateness for rare event analysis. Particularly, penalized Cox models were selected to address sparsity and multicollinearity issues, while the standard Cox model was chosen for its robustness in survival data analysis [19]. Meanwhile, for the machine learning approach, random survival forests and stochastic gradient boosting models were selected as they were deemed effective based on previous work done on renal graft failure datasets, as tree-based ensemble models are particularly effective for imbalanced data [20].



**Fig. 1** Computational workflow of the study

### Standard cox regression

It is a type of statistical model that is widely used to predict survival outcomes. The model estimates the hazard ratio for each variable, which represents the relative risk of the event occurring for a particular level of the variable. The model assumes a proportional hazards assumption, which means that the hazard ratio is constant over time [21]. Its mathematical expression is given by:

$$h(t | \mathbf{X}) = h_0(t) \exp \left( \sum_{i=1}^p \beta_i x_i \right) = h_0(t) \exp (\mathbf{T} \beta \mathbf{X}),$$

Where  $h(t | \mathbf{X})$  is the risk of graft failure given a set of predictors,  $h_0(t)$  is the baseline hazard (the hazard when all predictors are equal to zero),  $\beta = (\beta_1, \dots, \beta_p)^T$  is the column vector of regression parameters, and  $\mathbf{T} \beta$  means its transpose,  $\mathbf{X} = \mathbf{T} (x_1, \dots, x_p)$  denotes the set of considered predictor variables. After some algebraic manipulation, the log-partial likelihood function of the Cox model becomes:

$$l(\beta) = \sum_{r \in D} \left( \beta^T \mathbf{X}^{(r)} - \log \left( \sum_{j \in R} \exp \left( \beta^T \mathbf{X}^{(j)} \right) \right) \right).$$

The standard maximum likelihood estimation method can be applied to calculate unknown parameters ( $\beta$ ). As a general rule, logistic and Cox models can be fitted with at least 10 events of the outcome variable per predictor variable, to prevent overfitting [16]. However, in our study, the total number of outcome events (graft failure) is 21 out of 278 subjects and several predictor variables have been taken into account. As a result, there are relatively few events per predictor. This rare event results in overfitting and inconsistent estimates when fitted with the standard Cox regression. Penalization alleviates this issue by restricting the size of regression coefficients through the use of a complexity parameter that controls shrinkage [22]. The three commonly used Cox penalization models are the Lasso-based, Ridge-based, and Elastic net-based Cox models.

### Lasso-based cox regression

Lasso-based Cox regression is a type of Cox model that includes a penalty term to encourage sparsity in the model coefficients. This can help to identify the most important predictors and reduce overfitting. The model penalty term shrinks the coefficients towards zero, which can help to prevent the coefficients from becoming too large and unstable. To incorporate the regularization term into the Cox regression model, the log partial likelihood function would be rewritten as follows.

$$l(\beta) - \sum_{j=1}^p \lambda |\beta_j|,$$

where  $\lambda$  is a tuning (regularization) parameter and  $p$  is the number of predictors. The  $L_1$ -penalized (Lasso) characterized by simultaneous variable selection and shrinkage is a useful method for determining interpretable prediction rules in high-dimensional data [23].

### Ridge-based cox regression

Ridge-based Cox regression, on the other hand, is a regression method that shrinks the regression coefficients towards zero (not exactly zero) by imposing a penalty on the sum of the squared values of the coefficients. This technique is useful when there are many potentially relevant predictors for survival outcomes, as it can help stabilize the estimates of the coefficients in the presence of multicollinearity [24]. The  $L_2$ -penalized (ridge) log partial likelihood is written as:

$$l(\beta) - \sum_{j=1}^p \lambda \beta_j^2$$

### Elastic net-based cox regression

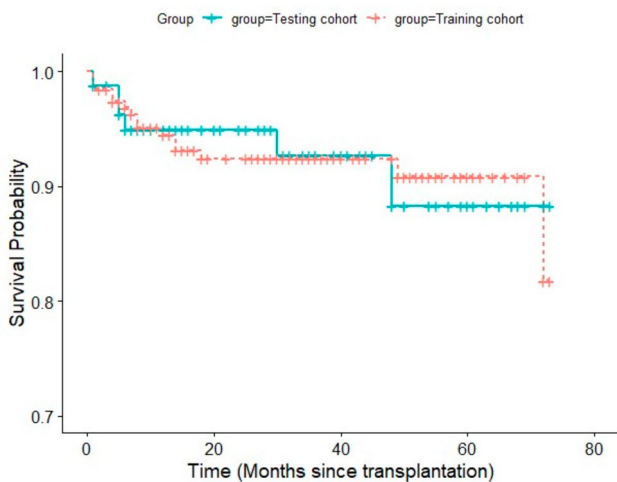
Elastic net-based Cox regression is a hybrid of lasso-based and ridge-based Cox regression. The model includes both  $L_1$  and  $L_2$  penalty terms, which can help to identify important predictors and prevent overfitting. The model is particularly useful when there are many correlated predictors and when the outcome is sparse [25]. Its penalization is given by:

$$l(\beta) - \sum_{j=1}^p (\lambda_1 |\beta_j| + \lambda_2 \beta_j^2),$$

where  $\lambda_1$  and  $\lambda_2$  are corresponding regularization parameters of  $L_1$  and  $L_2$  penalties, respectively.

### Random survival forest

In survival analysis, the main challenge of applying machine learning methods is the difficulty of appropriately dealing with censored information and the time estimation of the model [26]. Random survival forest is a type of ensemble machine learning model as an extension of random forest it uses a combination of decision trees and survival analysis techniques to predict the time-to-event outcome. The model randomly selects features and survival times to build a decision tree and then combines multiple decision trees to form a forest. Combining information from the separate trees like survival probabilities and ensemble cumulative hazard estimates can be calculated using the Kaplan-Meier and Nelson-Aalen methodologies, respectively. The model is useful for handling high-dimensional data, capturing non-linear relationships, and mitigating overfitting [27].



**Fig. 2** The Kaplan-Meier Survival Curves for Training and Testing Datasets

**Stochastic gradient boosting survival**

Stochastic gradient boosting survival is another machine learning model that uses an ensemble of decision trees to predict survival outcomes [28]. It is an extension of stochastic gradient boosting suited for survival predictions. The model iteratively adds decision trees to the ensemble, with each tree correcting the errors of the previous tree. The model can handle non-linear relationships between variables and can provide accurate and personalized predictions. The two machine learning algorithms (SGB and RSF) were selected based on their superior prediction performance in predicting renal graft failure [20].

**Model comparison**

Each model’s performance is assessed and compared based on its calibration, discrimination, and interpretability. We used global metrics like the concordance index (C-index) and Brier score, as well as graphical methods such as the ROC curve and calibration plot. In the evaluation process, we followed a 3-stage comparison approach: first, we assessed the model’s discrimination

(accuracy), followed by calibration (absolute accuracy), which is often neglected despite its importance (only 36% of published models provided a calibration measure). Finally, we evaluated the model’s interpretability, in terms of the most prognostic factors identified [29].

**Results**

**Patient characteristics**

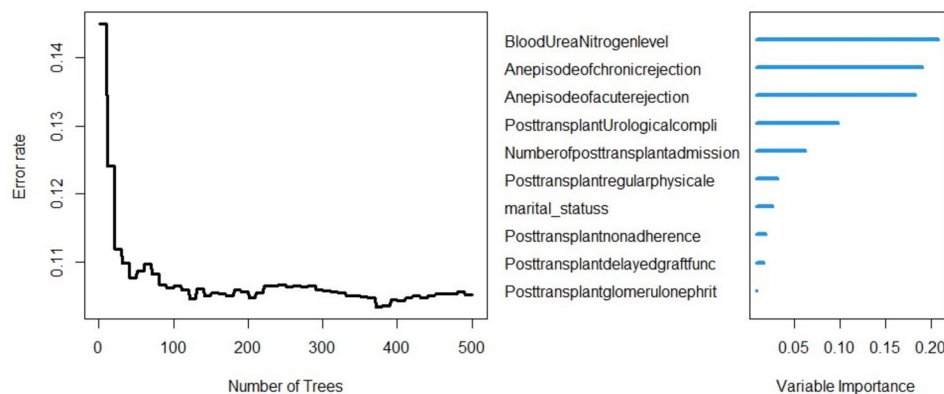
The study cohort consists of a total of 278 transplant recipients under the follow-up of the Ethiopian National Kidney Transplant Center. The survival times of the graft ranged from 1 to 73 months, with an average hazard of graft failure of 0.0755 and a median survival time of 33 months. The 3-month, 1-year, and 3-year graft survival rates were found to be 0.979, 0.953, and 0.911, respectively. Of the patients, 74.8% were male and 25.2% were female, with a median age of 37 years. The original entire group was divided into training and testing sets to develop and test clinical prognostic models. It was essential to ensure that the survival rates of both datasets were proportional. The Kaplan-Meier survival function and log-rank test were utilized to confirm this, and it was found that there was no significant difference ( $p = 0.96$ ) in survival probability between the training and testing sets, as shown in Fig. 2.

**Construction of the clinical prognostic models**

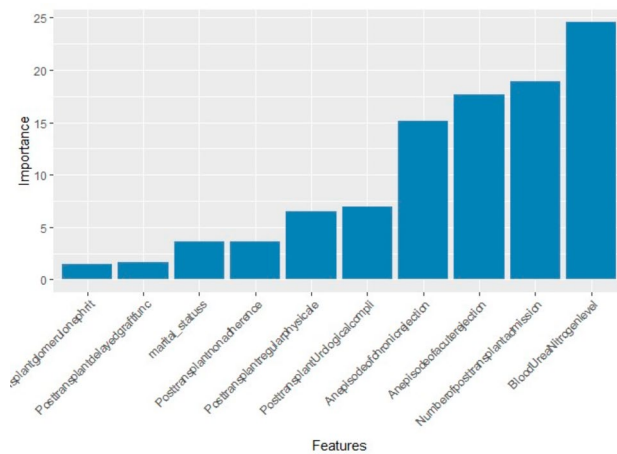
This study involved developing a range of statistical and machine-learning models to predict renal graft survival by utilizing a balanced dataset and carefully chosen features. After the models were developed with a cross-validation hyperparameter tuning technique, their predictive performance was evaluated using a testing dataset that remained unseen by the models and was not resampled.

**Random survival forests (RSF)**

The RSF model was developed using 5-fold cross validation and hyper parameter tuning. Important results of the model were presented in Fig. 3. Based on the figure,



**Fig. 3** a) Error rate across tree numbers; b) variable importance from random survival forests



**Fig. 4** Variable importance from the SGB model

the error rate was relatively low and started to stabilize when the number of survival trees was 500. The feature importance score of each prognostic variable was calculated and the features were ranked in descending order. Accordingly, blood urea nitrogen level, an episode of chronic rejection, an episode of acute rejection, post-transplant urological complications, the number of post-transplant admissions, post-transplant regular physical exercise, and marital status were the top seven prognostic features for renal graft survival.

**Stochastic gradient boosting (SGB) model**

Similarly, the hyper parameter tuning for SGB model was performed using 5-fold cross validation. As depicted in Fig. 4, the top seven predictors of graft survival are blood urea nitrogen level, the number of post-transplant admissions, an episode of acute rejection, an episode of chronic rejection, post-transplant urological complications, post-transplant regular physical exercise, and post-transplant nonadherence.

**The standard cox PH model**

The standard Cox regression model was developed using the Breslow estimation method. After fitting the model,

proportionality was tested and achieved with a global p-value of 0.650. According to the Cox PH regression model results in Table 1, an episode of acute rejection, higher blood urea nitrogen level, an episode of chronic rejection, post-transplant urological complication, and post-transplant non-adherence were associated with a significantly higher risk of graft failure ( $p < 0.05$ ), whereas, patients who are married (cohabited) and those who perform regular physical exercise after the transplant have a significantly lower risk of graft failure.

Based on the Table, patients who experienced an episode of acute rejection had a 3.782 times higher risk of graft failure compared to those who did not have an acute rejection episode. Similarly, every one-unit increase in blood urea nitrogen (BUN) level was associated with a 39.6% higher risk of graft failure. Patients who developed an episode of chronic rejection had a 2.136 times greater risk of graft failure relative to those without chronic rejection. The presence of a post-transplant urological complication was also a significant risk factor, increasing the risk of graft failure by 1.735 times. Additionally, patients who were non-adherent to the post-transplant treatment regimen had a 1.640 times higher risk of graft failure compared to adherent patients.

In contrast, the analysis identified two protective factors against graft failure. Being married (cohabited) was associated with a 40.8% lower risk of graft failure compared to living alone (single, divorced, separated, or widowed). This may be due to the emotional, practical, and financial support a partner can offer. This helps patients follow the complicated medication schedules after a transplant keep up with medical appointments, and cope with the recovery demands. Such a social support network can have a beneficial impact on mental health and overall well-being, which are vital for positive graft outcomes, like elongated graft survival. Moreover, engaging in regular physical exercise after the transplant was even more strongly protective, reducing the risk of graft failure by 73.8% compared to those who did not exercise regularly.

**Table 1** Cox Regression Analysis Results for predictors of renal graft survival

Variables	Coef.	Exp (Coef.)	Se (Coef.)	z	Pr.(> z )
An episode of acute rejection	1.330	3.782	0.292	4.561	<0.001
Blood Urea Nitrogen level	0.333	1.396	0.091	3.678	<0.001
An episode of chronic rejection	0.759	2.136	0.274	2.774	0.006
Post-transplant Urological complication	0.551	1.735	0.258	2.136	0.033
Post-transplant nonadherence	0.495	1.640	0.242	2.046	0.041
Post-transplant delayed graft functioning	0.347	1.415	0.247	1.404	0.160
Number of post-transplant admissions	0.061	1.063	0.090	0.680	0.496
Post-transplant glomerulonephritis	0.139	1.149	0.265	0.524	0.600
Marital status	-0.523	0.592	0.225	-2.323	0.020
Post-transplant regular physical exercise	-1.341	0.262	0.294	-4.561	<0.001

**Table 2** Lasso-Cox Regression Analysis results for predictors of renal graft survival

Variable	Coef.	Exp (Coef.)
An episode of acute rejection	1.247	3.480
An episode of chronic rejection	0.785	2.192
Post-transplant Urological complication	0.449	1.567
Post-transplant nonadherence	0.411	1.508
Blood Urea Nitrogen level	0.300	1.350
Post-transplant delayed graft functioning	0.230	1.258
Post-transplant glomerulonephritis	0.117	1.124
Number of post-transplant admissions	0.049	1.050
Marital status	-0.372	0.689
Post-transplant regular physical exercise	-1.032	0.356

**Table 3** Ridge-Cox Regression Analysis results for predictors of renal graft survival

Variable	Coef.	Exp (Coef.)
Blood Urea Nitrogen level	0.288	1.334
Number of post-transplant admissions	0.084	1.087
Post-transplant regular physical exercise	-1.015	0.362
Post-transplant nonadherence	0.469	1.598
Post-transplant Urological complications	0.533	1.705
An episode of acute rejection	1.154	3.171
An episode of chronic rejection	0.741	2.098
Post-transplant glomerulonephritis	0.156	1.169
Post-transplant delayed graft functioning	0.336	1.400
Marital status	-0.463	0.630

#### Lasso-based Cox model

The lasso-based Cox model is a type of Cox model that includes a penalty term to encourage sparsity in the model coefficients. In this case, all ten predictors were returned as non-zero coefficients, as the data-driven pre-feature selection was conducted before the model development phase. In the order of their effect size, acute rejection, chronic rejection, post-transplant regular physical exercise, post-transplant urological complications, post-transplant nonadherence, blood urea nitrogen level, and marital status were found to be the top seven significant prognostic predictors of graft survival, as presented in Table 2. Consistent with the result of the standard Cox model, regular physical exercise and marital status (being married or cohabited) appear to have a protective effect against the risk of graft failure.

#### Ridge-based Cox model

This is another regularized version of the Cox model that shrinks the regression coefficients towards zero (not exactly zero) by imposing a penalty on the sum of the squared values of the coefficients. Similarly, it returns ten features with non-zero coefficients, indicating that the coefficients were penalized but not reduced to zero. According to Table 3, the seven leading predictors of graft survival, ranked by effect size, are acute rejection, chronic rejection, post-transplant urological

**Table 4** EN-Cox Regression Analysis results for predictors of renal graft survival

Variable	Coef.	Exp (Coef.)
Blood Urea Nitrogen level	0.291	1.338
Number of post-transplant admissions	0.063	1.065
Post-transplant regular physical exercise	-0.998	0.369
Post-transplant nonadherence	0.423	1.526
Post-transplant Urological complications	0.474	1.606
An episode of acute rejection	1.200	3.321
An episode of chronic rejection	0.765	2.149
Post-transplant glomerulonephritis	0.125	1.133
Post-transplant delayed graft functioning	0.262	1.299
Marital status	-0.394	0.674

complications, post-transplant regular physical exercise, post-transplant nonadherence, post-transplant delayed graft functioning, and marital status.

#### Elastic Net-based Cox model (EN-based Cox)

The EN-based Cox model is a hybrid of lasso-based and ridge-based Cox regression models. Based on the EN-based Cox results in Table 4, acute rejection, chronic rejection, post-transplant regular physical exercise, post-transplant urological complications, post-transplant nonadherence, blood urea nitrogen levels, and marital status have been identified as the top seven significant prognostic predictors of graft survival, ranked by the magnitude of their effect size. Post-transplant regular physical exercise and marital status (being married or cohabited) still have a protective effect against the risk of graft failure.

#### Identification of significant predictors

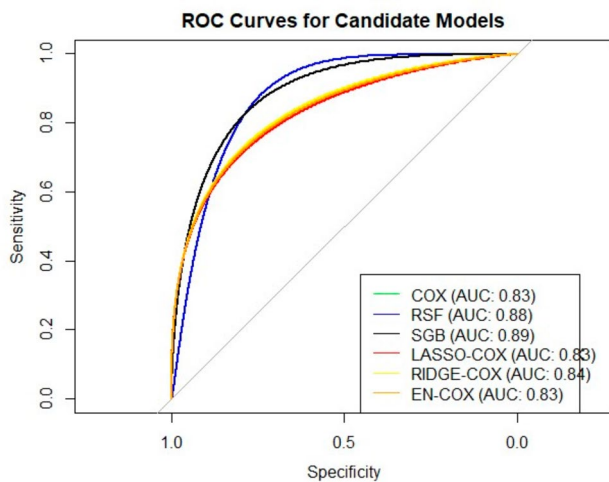
To determine the significant predictors of renal graft survival, we employed a combination of statistical and machine-learning models. The selection of predictor variables was based on their identification as the top seven significant predictors by most models, with particular emphasis on the best-performing and most interpretable models. Each predictor selected within the top seven was ranked from 1 to 7 across each model and marked with an 'x' if not selected (Table 5). Through a comprehensive ranking of predictors across the prognostic models, episodes of acute and chronic rejection, post-transplant regular physical exercise, post-transplant urological complications, post-transplant nonadherence, blood urea nitrogen levels, and marital status emerged as the most significant prognostic predictors of renal graft survival.

The findings suggest that regular physical exercise and being married or cohabiting are significantly associated with a reduced risk of graft failure. Consistent physical activity not only enhances overall health but also promotes better immune function and adherence to medical protocols. Likewise, having a supportive partner can

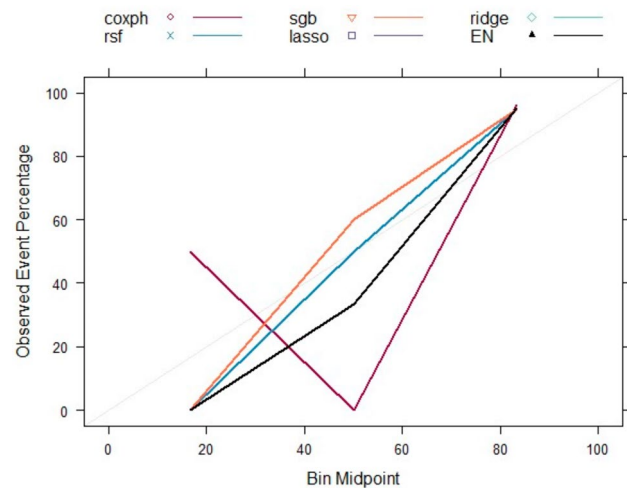


**Table 5** The variables chosen as the top seven predictors by at least one model

Variable	Cox.	Lasso-Cox	Ridge-Cox	EN-Cox	RSF	SGB
An episode of acute rejection	1	1	1	1	3	3
Blood Urea Nitrogen level	6	6	x	6	1	1
An episode of chronic rejection	2	2	2	2	2	4
Urological complications	4	4	3	4	4	5
Post-transplant nonadherence	5	5	5	5	x	7
Regular physical exercise	3	3	4	3	6	6
Marital status	7	7	7	7	7	x
Number of post-transplant admissions	x	x	x	x	5	2
Delayed graft functioning	x	x	6	x	x	x



**Fig. 5** The ROC curves for each candidate model



**Fig. 6** The calibration plot for each candidate model

improve psychological well-being, which is crucial for maintaining adherence to post-transplant care.

In contrast, several factors contribute to an increased risk of graft failure. Episodes of acute and chronic rejection are critical events that can severely compromise graft integrity and function. Additionally, urological complications may lead to further medical issues that jeopardize graft health. High blood urea nitrogen levels serve as vital indicators of kidney function; elevated levels often signify deteriorating graft health. Furthermore, post-transplant nonadherence to prescribed medication regimens significantly undermines the likelihood of long-term graft success. A more detailed discussion of these predictors is provided in the discussion section.

**Validation and comparison of the candidate models**

The prediction performance of each model was validated using the testing dataset. We compared different statistical and machine learning models using various prediction performance metrics and clinical relevance. The ROC curve, calibration plot, concordance index (C-index), brier score, and interpretability of the model were used as evaluation criteria to compare the prognostic models. The ROC curve was used to compare the discrimination

performance of the models. The result of the ROC curve in Fig. 5 indicated that the stochastic gradient boosting models had the highest discrimination performance with an AUC of 0.89, followed by the random survival forest with an AUC of 0.88. The ridge-based Cox model was ranked third in discrimination performance with an AUC of 0.84, while the standard, lasso-based, and EN-based Cox models shared the fourth-highest discrimination performance with an AUC of 0.83.

The calibration plot in Fig. 6 suggests that the random survival forest and stochastic gradient boosting models have superior calibration performance, as indicated by their proximity to the ideal 45-degree line. The overlaid black line represents the penalized versions of the Cox model (lasso, ridge, and EN-based Cox), which also demonstrate good calibration. In contrast, the standard Cox model's calibration lines are noticeably further away.

Based on the C-index in Table 6, the Stochastic Gradient Boosting (SGB) model emerged as the best performer, with a C-index of 0.943. The Ridge-based Cox model came in second with a C-index of 0.932. A C-index above 0.80 is generally considered adequate for clinical applications, and our results demonstrate that both the SGB and the Ridge-based Cox model meet this criterion.

**Table 6** Model comparison using global performance measures

Metrics	Cox PH	Ridge-Cox	Lasso-Cox	EN-Cox	RSF	SGB
C-index	0.930	<b>0.932</b>	0.929	0.929	0.925	<b>0.943</b>
Brier Score	0.0181	<b>0.000389</b>	0.000435	0.000431	0.012769	<b>0.000351</b>

These scores demonstrate the model's superior ability to rank individuals according to their predicted survival probabilities. In terms of the Brier score, which evaluates the accuracy of predicted survival probabilities, the SGB model maintained its lead with the lowest score of 0.000351, highlighting its remarkable precision in survival probability predictions. The Ridge-based Cox model closely follows SGB with a Brier score of 0.000389. This low score highlights its precision in providing reliable risk assessments, which is essential for clinical decision-making. Accurate predictions can help clinicians tailor treatment plans and improve patient outcomes by effectively stratifying risk among renal transplant recipients.

The Cox model and its penalized versions are widely utilized in survival analysis due to their high interpretability. These models provide hazard ratios, which help in understanding the impact of predictor variables on survival outcomes. In contrast, while Stochastic Gradient Boosting (SGB) achieves impressive predictive performance, its interpretability is limited compared to Cox-based models, as it does not offer explicit hazard ratios or easily interpretable coefficients. However, SGB can provide insights into variable importance through feature importance rankings. For those Cox-based models, we primarily focused on interpreting the standard Cox model. The penalized versions yielded results consistent with those of the standard model when fitted based on features selected from the penalized Cox models. Thus, our interpretation of significance and effect size was centered on the standard Cox model's hazard ratios.

Although the SGB model demonstrated superior calibration and discrimination performance, it lacks the interpretability of the Ridge-based Cox model, which is our second-best model. The Ridge-Cox model not only delivers robust performance but also offers clear insights into the impact of individual predictors on renal graft survival. This balance between predictive accuracy and interpretability is crucial in clinical settings, where understanding the rationale behind model predictions can significantly influence decision-making. When the primary goal is accurate prediction, the SGB model is preferable. Conversely, if interpretability is prioritized, we recommend the Ridge-Cox model, even if it entails a slight trade-off in predictive performance. This nuanced approach allows clinicians to select the most suitable model based on their specific needs and the context of their practice.

## Discussions

This study aimed to compare statistical and machine learning models for predicting renal graft survival and identify significant prognostic predictors. The main findings of the study have been discussed. In the evaluation of various models, the study found that the Stochastic Gradient Boosting model demonstrated the best calibration and discrimination performance. This finding is consistent with previous studies [20, 30]. Based on the global performance measures, the Ridge-based Cox model has the second-best calibration and discrimination performance. As a Cox-based model, it also offers the advantage of interpretability, allowing for a clear understanding of the impact of predictors on renal graft survival. This finding is supported by previous studies [31, 32] that confirmed the superiority of the Cox-based model over the tree-based models in terms of interpretability.

Therefore, the choice of model should consider the trade-off between prediction performance and interpretability. It is crucial for clinicians to carefully evaluate the strengths and limitations of each model and consider the context of their clinical practice when making a decision. If the primary concern is an accurate prediction (calibration and discrimination), the SGB model demonstrates the best prediction performance. On the other hand, if interpretability on the impact of predictors with the expense of a slight loss in accuracy the Cox (Ridge-Cox) model would be recommended.

Regarding prognostic predictors, our analysis identified several variables that were significant predictors of renal graft survival, including an episode of acute and chronic rejection, post-transplant urological complications, post-transplant nonadherence, blood urea nitrogen level, post-transplant regular exercise, and marital status. The discussion of each significant predictor is provided as follows.

Episodes of chronic rejection were significant predictors of graft survival. Chronic rejection occurs when the recipient's immune system gradually damages the transplanted kidney over time. This predictor suggests that the occurrence of chronic rejection significantly impacts renal graft survival. This finding agrees with previous studies [33, 34] which states that patients with chronic rejection were subjected to shorter graft survival. Managing immunosuppressive therapy and closely monitoring patients with a history of chronic rejection may be crucial for improving graft survival.

Episodes of acute rejection were found as significant predictors of graft survival. Acute rejection refers to an

immune response against the transplanted kidney shortly (days to months) after the transplantation. This predictor indicates that experiencing episodes of acute rejection can negatively impact graft survival. This result is in line with the literature [20, 35, 36]. Prompt diagnosis, effective immunosuppression, and close monitoring are crucial in managing acute rejection and improving graft outcomes.

Post-Transplant Urological Complications were also found to be significant predictors of renal graft survival. Urological complications following a kidney transplant include urinary tract infections, ureteral obstruction, and vesicoureteral reflux. This predictor suggests that the occurrence of post-transplant urological complications is associated with decreased graft survival. This finding is consistent with previous studies [37, 38]. Early detection, appropriate management, and preventive measures can help mitigate the impact of urological complications on graft outcomes.

Post-transplant nonadherence significantly affects graft survival. Nonadherence refers to a patient's failure to adhere to the prescribed medication regimen or follow recommended lifestyle modifications after transplantation. This predictor indicates that nonadherence is a significant risk factor for graft failure and is associated with poor graft survival. This finding is well supported by the literature [39–41]. Patient education, counseling, and support systems are crucial in promoting adherence to immunosuppressive medications and post-transplant care, thereby improving graft survival.

The study found that blood urea nitrogen (BUN) level is a significant predictor of renal graft survival. BUN is a well-known biomarker of kidney function. Elevated BUN levels indicate impaired kidney function and inadequate clearance of urea, which can be attributed to poor graft function, medication non-adherence, or the presence of comorbidities. This is supported by previous studies [42, 43]. Monitoring BUN levels in kidney transplant recipients allows for early detection of graft dysfunction and the need for intervention, such as medication adjustments, patient education, lifestyle modifications, and comorbidity management, thereby improving graft survival.

Post-transplant regular physical exercise has been shown to have a protective effect against the risk of graft failure. Physical exercise can help maintain a healthy body weight, improve cardiovascular health, and reduce the risk of metabolic complications that can contribute to graft failure. It can also boost the immune system and help prevent rejection of the transplanted kidney. This finding is consistent with a previous study [44]. Healthcare providers should encourage their patients to maintain a regular exercise routine for overall well-being,

which can significantly reduce the risk of renal graft failure in transplant recipients.

Being married or cohabitated has been identified as a factor that can reduce the risk of renal graft failure compared to patients who live alone. Married or cohabitated individuals often have a spouse or partner who can provide valuable support, such as helping with medication adherence, attending medical appointments, and assisting with daily self-care tasks. This support system can improve the transplant recipient's mental health, reduce their stress levels, and encourage healthier behaviors - all of which contribute to better outcomes for the transplanted kidney. This protective effect of being married (cohabitated) against the risk of graft failure is supported by a previous study [45].

#### **Clinical implication, strength, and limitation of the study**

This study has important clinical implications. The prognostic models developed can help clinicians make informed decisions about the management of transplant patients. Early identification of high-risk patients for graft failure allows for early interventions, such as intensifying immunosuppression or closer monitoring. The models can also be used to track changes in the predicted probability of graft survival over time. Furthermore, the best models can guide the development of personalized treatment plans by incorporating patient-specific factors. These clinically relevant models can also inform patients and their families about the expected outcomes of renal transplantation ultimately leading to optimized long-term outcomes for recipients. We recommend that healthcare providers incorporate these models into clinical routine practice to standardize risk assessment and improve long-term outcomes for renal transplant recipients.

The main strength of this study is the use of a time-to-event dataset, which is more suitable for modeling graft survival compared to assuming known event status for all subjects. The comprehensive comparison of statistical and machine learning models provides a deeper understanding of the strengths and limitations of each approach, helping to identify the most effective method for predicting renal graft survival. The study's rigorous evaluation using measures like calibration, discrimination, and interpretability enhances the reliability and credibility of the findings. The inclusion of advanced techniques in survival prediction, such as random survival forest and stochastic gradient boosting survival, adds novelty and expands the methodological landscape. The study also addressed potential issues of data imbalance and overfitting and incorporated relevant clinical and prognostic predictors to improve the clinical relevance of the models.

The study has some limitations. While these issues have been effectively addressed, using small sample sizes and resampling techniques to tackle class imbalance can still introduce biases and restrict the generalizability of the models developed. The lack of external validation in independent datasets makes it difficult to assess the models' performance in different clinical settings. Lastly, the limited interpretability of ensemble machine learning models poses a challenge for clinical applications. Future research should address these limitations by using a larger sample and conducting external validation to strengthen the validity and generalizability of the findings.

## Conclusions

This study compared various statistical and machine learning predictive models for renal graft survival. The study found that the Stochastic Gradient Boosting model had the best calibration and discrimination performance. Moreover, Cox-based models offer great interpretability with a comparable prediction performance. Clinicians should consider the trade-off between accuracy and interpretability when choosing a model. The significant prognostic factors for renal graft survival were an episode of acute and chronic rejections, post-transplant urological complications, post-transplant nonadherence, blood urea nitrogen level, post-transplant regular exercise, and marital status (alone or cohabited). Incorporating these findings into clinical practice can improve personalized medicine and long-term outcomes for kidney transplant recipients.

## Abbreviations

ESRD	End Stage Renal Disease
ML	Machine Learning
RSF	Random Survival Forests
ROC	Receiver Operating Characteristic
AUC	Area under the Curve
C-index	Concordance Index
SGB	Stochastic Gradient Boosting
EN	Elastic Net
HLA	Human Leukemia Antigen
PH	Proportional Hazard
SMOTE	Synthetic Minority Over Sampling
ENKTC	Ethiopian National Kidney Transplantation Center

## Acknowledgements

We greatly acknowledged the kidney transplant nurses and the institutional review board of St. Paul's Hospital Millennium Medical College for their support and cooperation during the data collection process.

## Author contributions

GM contributed to the development of the proposal, supervision of data collection, data analysis, manuscript preparation, and editing of the manuscript. TZ contributed to supervising the overall activities, reading and editing the manuscript based on his rich experience. AST participated in preparing the manuscript and editing and commenting on the manuscript. MBM and LHJ provide a professional review of the proposal, data extraction tool, and manuscript from a clinical perspective. Each author contributed to the conception and design, acquiring data and drafting it critically to ensure intellectual contact.

## Funding

This research received no specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Data availability

The findings of this study are supported by data obtained from the National Kidney Transplantation Center at St. Paul's Hospital Millennium Medical College. However, the data is subject to certain restrictions and is not publicly available. If you require access to the data, you can contact the authors and obtain permission from the National Kidney Transplantation Center at St. Paul's Hospital Millennium Medical College.

## Declarations

### Ethics approval and consent to Participate

All experimental protocols in this study were conducted in accordance with the Declaration of Helsinki. Informed consent was obtained from all participants, including both donors and recipients. Ethical approval was granted by the St. Paul's Hospital Millennium Medical College Institutional Review Board (Reference Number: PM23/459) and the Bahir Dar University Ethical Review Committee (Reference Number: PRCSD/290/2014). The authors confirm that no organs or tissues were obtained from prisoners. All organs were sourced from the Ethiopian National Kidney Transplant Center, which operates under St. Paul's Hospital Millennium Medical College.

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

### Author details

<sup>1</sup>Department of Statistics, Bahir Dar University, Bahir Dar, Ethiopia

<sup>2</sup>School of Mathematics, Statistics & Computer Science, KwaZulu Natal University, Durban, South Africa

<sup>3</sup>Kidney Transplant Center, St. Paul's Hospital Millennium Medical College, Addis Ababa, Ethiopia

Received: 11 August 2024 / Accepted: 30 January 2025

Published online: 03 February 2025

## References

1. Al-Bahri S, et al. Bariatric surgery as a bridge to renal transplantation in patients with end-stage renal disease. *Obes Surg.* 2017;27:2951–5.
2. Lee Y, Bang, Kim DJ. How to establish clinical prediction models. *Endocrinol Metabolism.* 2016;31(1):38–44.
3. Topuz K, et al. Predicting graft survival among kidney transplant recipients: a bayesian decision support model. *Decis Support Syst.* 2018;106:97–109.
4. Loupy A et al. Prediction system for risk of allograft loss in patients receiving kidney transplants: international derivation and validation study. *BMJ.* 2019. 366.
5. Lee YH, et al. Advanced tertiary lymphoid tissues in protocol biopsies are associated with progressive graft dysfunction in kidney transplant recipients. *J Am Soc Nephrol.* 2022;33(1):186–200.
6. Ramspek CL, et al. External validation of prognostic models: what, why, how, when and where? *Clin Kidney J.* 2021;14(1):49–58.
7. Kantidakis G, et al. Survival prediction models since liver transplantation-comparisons between Cox models and machine learning techniques. *BMC Med Res Methodol.* 2020;20:1–14.
8. Bakas S et al. *Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge.* arXiv preprint arXiv:1811.02629, 2018.
9. Nemati M, Ansary J, Nemati N. Machine-learning approaches in COVID-19 survival analysis and discharge-time likelihood prediction using clinical data. *Patterns.* 2020. 1(5).
10. Cho SM, et al. Machine learning compared with conventional statistical models for predicting myocardial infarction readmission and mortality: a systematic review. *Can J Cardiol.* 2021;37(8):1207–14.

11. Smith H, et al. A scoping methodological review of simulation studies comparing statistical and machine learning approaches to risk prediction for time-to-event data. *Diagn Prognostic Res.* 2022;6(1):1–15.
12. Naqvi SAA, et al. Predicting kidney graft survival using machine learning methods: prediction model development and feature significance analysis study. *J Med Internet Res.* 2021;23(8):e26843.
13. Adler AI, Painsky A. Feature importance in gradient boosting trees with cross-validation feature selection. *Entropy.* 2022;24(5):687.
14. Bischl B, et al. Hyperparameter optimization: foundations, algorithms, best practices, and open challenges. *Wiley Interdisciplinary Reviews: Data Min Knowl Discovery.* 2023;13(2):e1484.
15. Jardillier R, et al. Prognosis of lasso-like penalized Cox models with tumor profiling improves prediction over clinical data alone and benefits from bi-dimensional pre-screening. *BMC Cancer.* 2022;22(1):1045.
16. Elreedy D, Atiya AF. A comprehensive analysis of synthetic minority oversampling technique (SMOTE) for handling class imbalance. *Inf Sci.* 2019;505:32–64.
17. Velarde G, et al. Tree boosting methods for balanced and imbalanced classification and their robustness over time in risk assessment. *Intell Syst Appl.* 2024;22:200354.
18. Moons KG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med.* 2015;162(1):W1–73.
19. Pavlou M et al. How to develop a more accurate risk prediction model when there are few events. *BMJ.* 2015. 351.
20. Mulugeta G, et al. Classification of imbalanced data using machine learning algorithms to predict the risk of renal graft failures in Ethiopia. *BMC Med Inf Decis Mak.* 2023;23(1):1–17.
21. Lee S, Lim H. Review of statistical methods for survival analysis using genomic data. Volume 17. *Genomics & informatics*; 2019. 4.
22. Suchting R, et al. Using elastic net penalized cox proportional hazards regression to identify predictors of imminent smoking lapse. *Nicotine Tob Res.* 2019;21(2):173–9.
23. Qian J, et al. A fast and scalable framework for large-scale and ultrahigh-dimensional sparse regression with application to the UK Biobank. *PLoS Genet.* 2020;16(10):e1009141.
24. van de Wiel MA, van Nee MM, Rauschenberger A. Fast cross-validation for multi-penalty high-dimensional ridge regression. *J Comput Graphical Stat.* 2021;30(4):835–47.
25. Gong C, et al. Elastic net-based identification of GAMT as potential diagnostic marker for early-stage gastric cancer. *Biochem Biophys Res Commun.* 2022;591:7–12.
26. Wang P, Li Y, Reddy CK. Machine learning for survival analysis: a survey. *ACM Comput Surv (CSUR).* 2019;51(6):1–36.
27. Pölsterl S, et al. Survival analysis for high-dimensional, heterogeneous medical data: exploring feature extraction as an alternative to feature selection. *Artif Intell Med.* 2016;72:1–11.
28. Xia Y, et al. A dynamic credit scoring model based on survival gradient boosting decision tree approach. *Technological Economic Dev Econ.* 2021;27(1):96–119.
29. Chang W-J, et al. Evaluating methodological quality of prognostic prediction models on patient reported outcome measurements after total hip replacement and total knee replacement surgery: a systematic review protocol. *Syst Reviews.* 2022;11(1):1–8.
30. Karhade AV, et al. Development of machine learning algorithms for prediction of mortality in spinal epidural abscess. *Spine J.* 2019;19(12):1950–9.
31. Qiu X, et al. A comparison study of machine learning (random survival forest) and classic statistic (cox proportional hazards) for predicting progression in high-grade glioma after proton and carbon ion radiotherapy. *Front Oncol.* 2020;10:551420.
32. Du M, et al. Comparison of the tree-based machine learning algorithms to Cox regression in predicting the survival of oral and pharyngeal cancers: analyses based on SEER database. *Cancers.* 2020;12(10):2802.
33. Scheffner I, et al. Patient survival after kidney transplantation: important role of graft-sustaining factors as determined by predictive modeling using random survival forest analysis. *Transplantation.* 2020;104(5):1095–107.
34. Waiser J, et al. Predictors of graft survival at diagnosis of antibody-mediated renal allograft rejection: a retrospective single-center cohort study. *Transpl Int.* 2020;33(2):149–60.
35. Martinez-Mier G, et al. Acute rejection is a strong negative predictor of graft survival in living-donor pediatric renal transplant: 10-year follow-up in a single Mexican center. *Exp Clin Transpl.* 2019;17(2):170–6.
36. Koo EH, et al. The impact of early and late acute rejection on graft survival in renal transplantation. *Kidney Res Clin Pract.* 2015;34(3):160–4.
37. Buttigieg J, et al. Early urological complications after kidney transplantation: an overview. *World J Transplantation.* 2018;8(5):142.
38. Friedersdorff F, et al. Long-term follow-up after paediatric kidney transplantation and influence factors on graft survival: a single-centre experience of 16 years. *Urol Int.* 2018;100(3):317–21.
39. Ndemera H, Bhengu B. Factors contributing to kidney allograft loss and associated consequences among post kidney transplantation patients. *Health Sci J.* 2017;11(3):1.
40. Gaynor JJ, et al. Graft failure due to nonadherence among 150 prospectively-followed kidney transplant recipients at 18 years post-transplant: our results and review of the literature. *J Clin Med.* 2022;11(5):1334.
41. Mohamed M, et al. Non-adherence to appointments is a strong predictor of medication non-adherence and outcomes in kidney transplant recipients. *Am J Med Sci.* 2021;362(4):381–6.
42. Lu H-Y, et al. Predictive value of serum creatinine, blood urea nitrogen, uric acid, and  $\beta$  2-microglobulin in the evaluation of acute kidney injury after orthotopic liver transplantation. *Chin Med J.* 2018;131(09):1059–66.
43. Kim D-G, et al. Quantitative ultrasound for non-invasive evaluation of sub-clinical rejection in renal transplantation. *Eur Radiol.* 2023;33(4):2367–77.
44. Ponticelli C, Favi E. Physical inactivity: a modifiable risk factor for morbidity and mortality in kidney transplantation. *J Personalized Med.* 2021;11(9):927.
45. Prihodova L, et al. Adherence in patients in the first year after kidney transplantation and its impact on graft loss and mortality: a cross-sectional and prospective study. *J Adv Nurs.* 2014;70(12):2871–83.

## Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.