ORIGINAL ARTICLE

Retrospective Cohort Study

Predicting major adverse cardiovascular events after orthotopic liver transplantation using a supervised machine learning model: A cohort study

Jonathan Soldera, Leandro Luis Corso, Matheus Machado Rech, Vinícius Remus Ballotin, Lucas Goldmann Bigarella, Fernanda Tomé, Nathalia Moraes, Rafael Sartori Balbinot, Santiago Rodriguez, Ajacio Bandeira de Mello Brandão, Bruno Hochhegger

Jonathan Soldera, Post Graduate Program at Acute Medicine and Gastroenterology, University of South Wales, Cardiff CF37 1DL, United Kingdom

Jonathan Soldera, Bruno Hochhegger, Postgraduate Program in Pathology, Federal University of Health Sciences of Porto Alegre (UFCSPA), Porto Alegre 90050-170, Brazil

Leandro Luis Corso, Fernanda Tomé, Nathalia Moraes, Department of Engineering, Universidade de Caxias do Sul, Caxias do Sul 95070-560, Brazil

Matheus Machado Rech, Vinícius Remus Ballotin, Lucas Goldmann Bigarella, Rafael Sartori Balbinot, School of Medicine, Universidade de Caxias do Sul, Caxias do Sul 95070-560, Brazil

Santiago Rodriguez, Ajacio Bandeira de Mello Brandão, Postgraduate Program in Hepatology, Federal University of Health Sciences of Porto Alegre (UFCSPA), Porto Alegre 90050-170, Brazil

Corresponding author: Jonathan Soldera, MD, PhD, Instructor, Post Graduate Program at Acute Medicine and Gastroenterology, University of South Wales, Llantwit Rd, Pontypridd, Cardiff CF37 1DL, United Kingdom. jonathansoldera@gmail.com

Abstract

BACKGROUND

Liver transplant (LT) patients have become older and sicker. The rate of post-LT major adverse cardiovascular events (MACE) has increased, and this in turn raises 30-d post-LT mortality. Noninvasive cardiac stress testing loses accuracy when applied to pre-LT cirrhotic patients.

AIM

To assess the feasibility and accuracy of a machine learning model used to predict post-LT MACE in a regional cohort.

METHODS

This retrospective cohort study involved 575 LT patients from a Southern Brazilian academic center. We developed a predictive model for post-LT MACE

(defined as a composite outcome of stroke, new-onset heart failure, severe arrhythmia, and myocardial infarction) using the extreme gradient boosting (XGBoost) machine learning model. We addressed missing data (below 20%) for relevant variables using the k-nearest neighbor imputation method, calculating the mean from the ten nearest neighbors for each case. The modeling dataset included 83 features, encompassing patient and laboratory data, cirrhosis complications, and pre-LT cardiac assessments. Model performance was assessed using the area under the receiver operating characteristic curve (AUROC). We also employed Shapley additive explanations (SHAP) to interpret feature impacts. The dataset was split into training (75%) and testing (25%) sets. Calibration was evaluated using the Brier score. We followed Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis guidelines for reporting. Scikit-learn and SHAP in Python 3 were used for all analyses. The supplementary material includes code for model development and a user-friendly online MACE prediction calculator.

RESULTS

Of the 537 included patients, 23 (4.46%) developed in-hospital MACE, with a mean age at transplantation of 52.9 years. The majority, 66.1%, were male. The XGBoost model achieved an impressive AUROC of 0.89 during the training stage. This model exhibited accuracy, precision, recall, and F1-score values of 0.84, 0.85, 0.80, and 0.79, respectively. Calibration, as assessed by the Brier score, indicated excellent model calibration with a score of 0.07. Furthermore, SHAP values highlighted the significance of certain variables in predicting postoperative MACE, with negative noninvasive cardiac stress testing, use of nonselective beta-blockers, direct bilirubin levels, blood type O, and dynamic alterations on myocardial perfusion scintigraphy being the most influential factors at the cohort-wide level. These results highlight the predictive capability of our XGBoost model in assessing the risk of post-LT MACE, making it a valuable tool for clinical practice.

CONCLUSION

Our study successfully assessed the feasibility and accuracy of the XGBoost machine learning model in predicting post-LT MACE, using both cardiovascular and hepatic variables. The model demonstrated impressive performance, aligning with literature findings, and exhibited excellent calibration. Notably, our cautious approach to prevent overfitting and data leakage suggests the stability of results when applied to prospective data, reinforcing the model's value as a reliable tool for predicting post-LT MACE in clinical practice.

Key Words: Liver transplantation; Major adverse cardiac events; Machine learning; Myocardial perfusion imaging; Stress test

Core Tip: This study presents a robust machine learning model, utilizing the XGBoost algorithm, to predict major adverse cardiovascular events (MACE) following liver transplantation. The model demonstrated high accuracy and calibration, with key factors such as noninvasive cardiac stress test outcomes, use of nonselective beta-blockers, direct bilirubin levels, blood type O, and dynamic alterations on myocardial perfusion scintigraphy identified as significant predictors. This tool offers valuable insights into the risk assessment of post-liver transplant MACE, particularly in an aging and comorbid patient population.

INTRODUCTION

The population of liver transplant (LT) candidates has become older and sicker, experiencing higher morbidity[1]. This might be due to the increasing prevalence of metabolic-associated fatty liver disease (MAFLD) as a cause of cirrhosis and end-stage liver disease (ESLD)[2-5]. As a result, there is an expected rise in the incidence of major adverse cardiovascular events (MACE) following LT, a well-documented complication of LT that negatively impacts prognosis[6-10].

The occurrence of MACE in the post-LT period is a significant concern, since these events contribute to increased mortality and jeopardize the success of LT[11]. Previous literature suggests that the incidence of post-LT MACE can be as high as 41% within the first 6 months following LT, which translates into a higher mortality rate[6,10]. Various traditional and nontraditional cardiovascular risk factors may contribute to these adverse events, including preexisting coronary disease, obesity, reduced cardiovascular reserve, poor response to cardiovascular stress, cirrhotic cardiomyopathy, increased predisposition to arrhythmias, and heart failure exacerbations[12-15]. The prioritization for transplant of sicker

patients with a high burden of critical illness, associated with a higher prevalence of cardiovascular disease, further exacerbates the risk[16]. However, the relative contribution of these factors remains incompletely characterized[7,17,18].

In addition to population aging, there has been a significant change in the most prevalent etiology leading to LT, with an increase in MAFLD observed both in the West and in the East[2,19]. Currently, MAFLD is the fastest-growing indication for LT in Western countries, having become the leading indication for LT waitlisting in the United States[5], as predicted by previous studies[20]. Moreover, MAFLD is strongly associated with a higher prevalence of diabetes mellitus, morbid obesity, and coronary artery disease (CAD)[4,5,8]. This specific population thus requires a detailed pre-LT cardiac evaluation, with particular attention to the increased risk of CAD, as they have a higher risk of cardiac events compared to those without MAFLD[8,21].

The first stage of cardiac evaluation usually involves assessing risk factors and subsequently performing noninvasive stratification. However, this approach is still controversial. In 2014, the American Association for the Study of Liver Diseases updated its guideline, maintaining the recommendation that patients undergoing pre-LT evaluation should complete a noninvasive myocardial stress test[22]. Conversely, the 2012 guideline developed by the American Heart Association in conjunction with the American College of Cardiology[23] suggests performing a noninvasive myocardial stress test only for patients with three or more risk factors for CAD. However, systematic reviews have demonstrated that current noninvasive strategies, such as myocardial perfusion scintigraphy (MPS) and dobutamine stress echocardiography (DSE), are unreliable and inadequate for predicting MACE, mortality, and significant CAD after LT[24-26]. Therefore, there is an unmet need for an alternative approach to accurately predict post-LT MACE in this vulnerable patient population[18,27].

Few models are available to assist clinicians in accurately stratifying the cardiovascular risk of LT candidates, especially those with ESLD[18]. Existing models often rely on traditional logistic regression statistics, making assumptions of independent linear relationships between dependent and independent variables[28]. These models are further constrained by small sample sizes and the limited number of variables for which they can account, primarily due to concerns of overfitting and multicollinearity. They are also unable to accurately consider the small effects of minor variables and their complex correlations[18,28]. Two scores have been developed using such models, the CAD-LT[29], and the CAR-orthotopic liver transplantation (OLT)[30]. The CAD-LT has demonstrated ability to stratify the risk of CAD into low, intermediate, and high categories, while the CAR-OLT point-based prediction model has shown superior performance compared to other existing risk models in predicting post-LT MACE.

In addition, patients with liver cirrhosis exhibit significant peripheral vasodilation, which can alter cardiac function and mask the presence of CAD, leading to what is now termed cirrhotic cardiomyopathy, a distinct pathologic entity for which diagnostic criteria were published in 2020[31]. In the 1990s, a high mortality rate (around 50%) was reported in patients with significant CAD in the peri-LT period[32]. However, in the last decade, with improved pre-LT cardiac therapy, it is believed that the presence of CAD does not significantly alter the post-LT survival of these patients[33].

To overcome these limitations, we propose the use of machine learning, a subarea of computer science that focuses on predicting outcomes using computational models that iteratively learn from data[34,35]. Machine learning models have demonstrated robust performance in various fields in gastroenterology[36], such as the diagnosis of hepatocellular carcinoma[37], prognostication of variceal hemorrhage[38,39], prediction of acute kidney injury after LT[40], short- and long-term post-LT mortality[41], and adverse cardiovascular events in various medical conditions[42]. Unlike conventional statistical models, machine learning models can detect complex patterns and relationships within datasets without relying on fixed assumptions about data behavior or pre-selection of variables, using correlations within variables to determine outcome[43].

The aim of this study is to conduct a comprehensive assessment of the feasibility and accuracy of employing a machine learning model for prediction of MACE following LT. The study focuses on a specific regional cohort to examine the potential of machine learning techniques in effectively forecasting post-LT MACE. By leveraging advanced computational models, this research aims to enhance the predictive capabilities in identifying individuals at higher risk of experiencing MACE after LT, thereby enabling early intervention strategies and optimizing patient care.

MATERIALS AND METHODS

This retrospective cohort study was approved by the Research Ethics Committee of Universidade Federal de Ciências da Saúde de Porto Alegre under protocol no. 07793412.2.3001.5345 on May 22, 2013, and conducted in accordance with the ethical guidelines of the 1975 Declaration of Helsinki. The study utilized medical records from Irmandade Santa Casa de Misericórdia de Porto Alegre (Rio Grande do Sul, Brazil).

Inclusion and exclusion criteria
Patients above 18 years of age who underwent their first LT at Irmandade Santa Casa de Misericórdia de Porto Alegre, Guido Cantisani LT Team, Brazil, for cirrhosis, between January 1, 2001, and December 31, 2011, were eligible. Patients without cirrhosis, those with incomplete medical records, those who did not undergo cardiac evaluation prior to LT, retransplantation cases, and living-donor LT recipients were excluded. Patients with 20% or more missing data were excluded.

Outcomes
Data were systematically collected on structured forms encompassing extensive clinical and laboratory variables from the pre-LT, perioperative, and post-LT periods. The primary outcome of interest was any in-hospital MACE, a composite

outcome including stroke, new-onset heart failure, severe arrhythmia, and myocardial infarction. Statistics, including frequency, means, SD, and tests such as Pearson's χ^2 test and linear model analysis of variance (ANOVA), were conducted in R software (version 4.3.2) using the 'readxl' and 'dplyr' packages, with the analysis involving data manipulation and exploration.

Machine learning approach and model definition

We employed the extreme gradient boosting (XGBoost) model, available through the XGBoost package, to construct a classification model aimed at predicting post-LT MACE. XGBoost is particularly effective in handling imbalanced datasets and offers native support for missing data and categorical variables, making it particularly useful for real-world applications. The columns considered to compose the outcome variable were not included in the model to avoid bias and collinearity.

Data pre-processing and feature engineering

The dataset was divided into training (75%) and test (25%) sets, preserving the outcome proportions in both subsets[44]. The training set is used to teach the model, and the test set is used to evaluate how well the model has learned. To mitigate the risk of introducing bias by excluding patients with missing values, we employed a two-step imputation process using the Scikit-Learn package. First, we removed variables that had missing values for more than 20% of the patient population. Following this, we used the k-nearest neighbor (kNN) imputation algorithm to fill in the missing values for the remaining continuous variables, imputing the calculated mean value among the 10 closest neighbors. Of 83 features screened, the model incorporated 50 according to the measure of the impact of each feature on the model's prediction for an instance. This included patient demographics, laboratory data, medical history, and pre-LT cardiac evaluations, selected after an initial screening. Categorical and numerical variables were imputed using mode and kNN imputation, respectively. To avoid data leakage, transformations were first trained on the training dataset, and only then applied to test data. To simulate real-world settings in which missing data are often present, we trained an additional model without the imputation and one-hot step and describe its results following the main model report.

Model training and hyperparameter optimization

Overfitting is a problem that occurs when a machine learning model learns the training data too well and is unable to generalize to new data. This can happen when the model is too complex or when the training dataset is too small or noisy. As a result, the model outputs extremely accurate results in the training set but performs poorly on unseen test-set data. To avoid overfitting, we applied regularization and early-stop techniques during the training of the model, as described in the code. Regularization is a technique that penalizes the model for being too complex; early stopping is a technique that stops training the model when it starts to overfit the training data.

Hyperparameters are external configurations for the model that are not learned from the data and are used to optimize the model's performance. The training set was used for model training, while the test set was reserved for performance evaluation. The Optuna package was used for hyperparameter optimization. Additional information about the model hyperparameter results and training are provided as supplemental material.

Performance assessment

The area under the receiver operating characteristic curve (AUROC) was used as an evaluation metric and reported with a 95% confidence interval (CI). To calculate the AUROC, the true positive rates are compared against the false positive rates at various threshold settings. The AUROC represents the degree or measure of separability, indicating how well the model distinguishes between the classes.

The model's performance in predicting positive cases was also assessed using the area under the precision-recall curve (AUC-PR). The AUC-PR is a graphical representation of a model's precision and recall at different thresholds, which are the points where the model decides which class an instance belongs to. It is particularly useful when the classes are imbalanced. The x-axis represents recall (the proportion of actual positive cases that were correctly classified) and the y-axis represents precision (the proportion of cases classified as positive that are indeed positive). A higher AUC-PR indicates better performance in distinguishing between the classes.

In evaluating the model's ability to predict positive cases, additional metrics were employed, such as recall, precision, sensitivity, specificity, accuracy, and F1-score. Recall measures the model's effectiveness in correctly identifying actual positive cases among all positive instances. It is calculated by dividing the number of true positives by the sum of true positives and false negatives. Precision assesses the accuracy of the model's positive predictions by calculating the proportion of true positives among all instances predicted as positive, determined by dividing the number of true positives by the sum of true positives and false positives. Sensitivity evaluates the model's capability to identify positive cases accurately, similar to recall. Specificity measures the model's ability to correctly identify negative cases by calculating the proportion of true negatives among actual negatives. Accuracy reflects the overall correctness of the model's predictions, considering both true positives and true negatives relative to the total number of predictions. F1-score represents the harmonic mean of precision and recall, providing a balanced assessment of the model's performance. The statistical methods of this study were reviewed by co-author Corso LL.

Calibration assessment

Calibration is the process of refining the model to ensure that the predicted probabilities of an event occurring align well with the actual probabilities. We tested various methods of calibration for the validation model, including sigmoid, isotonic, and Gaussian calibration. We used calibration curves to present the comparison graphically. We used the Brier

score to choose the model with the best calibration for deployment and explanation of feature importance.

Model explanation and interpretation

The Shapley additive explanations (SHAP) framework was used to interpret the output of machine learning models, providing a measure of the impact of each feature on the model's prediction for an instance. SHAP are based on game theory and assign an importance value to each feature in a model. Features with positive SHAP values positively impact the prediction, while those with negative values have a negative impact. The magnitude of the SHAP value is a measure of how strong the effect is. To calculate SHAP values, we consider all possible combinations of features (coalitions) and how they affect the model's prediction. We then average the marginal contribution of each feature across all possible coalitions. This gives us a measure of how much each feature contributes to the model's prediction, taking into account the interactions between features.

Checklist adherence

In accordance with the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement, we have followed a comprehensive reporting framework for this study. The TRIPOD statement guided the design, implementation, and reporting of our prediction model for post-LT MACE and the respective checklist for present study is presented as a Supplementary material. The checklist comprised the 22 items outlined in the TRIPOD statement, ensuring transparency and rigor in our methodology and reporting.

Code availability and web deployment

The code employed for data preprocessing, feature engineering, and model development and evaluation is in an accessible public repository (link provided in the supplemental materials). Furthermore, we have deployed our model as a user-friendly MACE prediction calculator, which is now available online at https://huggingface.co/spaces/mmrech/mace-calc. The frontend application was coded with the Streamlit library. The model was originally saved and then loaded as a joblib file, and the backend application was deployed with Hugging Face Spaces. All phases from data preprocessing to model deployment were implemented in Python 3.

RESULTS

A comprehensive search of hospital databases identified a total of 662 patients who had undergone LT during the study period. From this initial cohort, 82 patients were excluded based on specific criteria. The reasons for exclusion were as follows: 19 patients transplanted due to fulminant liver failure, 32 patients who had undergone retransplantation, 7 patients transplanted due to familial amyloid polyneuropathy, 1 patient excluded due to amyloidosis without cirrhosis, 1 patient due to congenital hepatic fibrosis, 27 patients due to insufficient cardiological data, 2 patients who received living-donor grafts, 2 patients with primary hyperoxaluria, 2 patients with polycystic liver disease, and 1 patient with metastasis of a neuroendocrine tumor. Another 38 patients were excluded due to the high rate of missing data among selected variables. The dataset utilized by the final model consisted of 537 samples, with 23 events and 514 non-events (Figure 1). As noted above, the original dataset was split such that 75% was used for training the model and 25% was reserved as unseen data for internal validation. The proportion of outcomes (4.46%) was maintained in both the training and the validation sets.

General cohort

Of the 537 included patients, 23 developed in-hospital MACE, with a mean age at transplantation of 52.9 years. The majority, 66.1%, were male. The overall incidence of the composite variable MACE was 4.46%. The components of this outcome - stroke, new-onset heart failure, severe arrhythmia, and myocardial infarction - had observed rates of 0.19%, 1.3%, 1.3% and 1.67%, respectively. Detailed data on the general population included, the 50 variables used in model construction, and the composite outcomes are available in Table 1, specifying values for the total cohort, for the strata of present and absent MACE, and also their respective missing rates.

Model performance

The XGBoost model demonstrated substantial predictive capability, with an AUROC of 0.89. The classification results showed a precision of 0.89, recall of 0.80, and F1-score of 0.84 for the negative class. The AUROC and AUC-PR, along with their respective 95%CIs, are provided in Figure 2. The hyperparameters utilized for the best-performing model after optimization are provided in the supplementary materials, as is an overview regarding the role of these components in the model functionality.

Calibration

The model achieved optimal calibration with the isotonic method, as evidenced by the lowest Brier score of 0.100. This calibration demonstrated a high level of precision, recall, F1-score, and accuracy for both negative and positive classes, with closer proximity to the diagonal line on the calibration curve (Supplementary Figure 1). Calibration curve is provided as a supplemental material.

Table 1 Cohort patient data

Variable	Total (*n* = 537)	MACE absent (*n* = 520)	MACE present (*n* = 17)	*P* value
Race[1]				0.41
White	463 (90.4%)	447 (90.1%)	16 (100.0%)	
Mixed/other	28 (5.6%)	28 (5.6%)	0 (0.0%)	
Black	21 (4.1%)	21 (4.2%)	0 (0.0%)	
Missing	25	24	1	
Sex[1]				0.26
Male	352 (65.7%)	343 (66.1%)	9 (52.9%)	
Female	184 (34.3%)	176 (33.9%)	8 (47.1%)	
Missing	1	1	0 (0)	
Previous esophageal variceal ligation[1]				0.71
No	327 (64.5%)	318 (64.6%)	9 (60.0%)	
Yes	180 (35.5%)	174 (35.4%)	6 (40.0%)	
Missing	30	28	2	
Portal hypertensive gastropathy[1]				0.48
Mild	229 (49.4%)	220 (49.0%)	9 (60.0%)	
Absent	83 (17.9%)	82 (18.3%)	1 (6.7%)	
Intense	152 (32.8%)	147 (32.7%)	5 (33.3%)	
Missing	73	71	2	
Previous ascites[1]				0.05
Yes	393 (74.3%)	377 (73.6%)	16 (94.1%)	
No	136 (25.7%)	135 (26.4%)	1 (5.9%)	
Missing	8	8	0	
Previous spontaneous bacterial peritonitis[1]				0.92
No	385 (74.3%)	374 (74.4%)	11 (73.3%)	
Yes	133 (25.7%)	129 (25.6%)	4 (26.7%)	
Missing	19	17	2	
Previous hepatopulmonary syndrome[1]				0.31
No	399 (77.3%)	389 (77.6%)	10 (66.7%)	
Yes	117 (22.7%)	112 (22.4%)	5 (33.3%)	
Missing	21	19	2	
Previous use of nonselective beta-blockers[1]				0.01
No	270 (53.9%)	267 (54.8%)	3 (21.4%)	
Yes	231 (46.1%)	220 (45.2%)	11 (78.6%)	
Missing	36	33	3	
Portal vein thrombosis[1]				0.04
No	443 (85.7%)	432 (86.2%)	11 (68.8%)	
Yes	74 (14.3%)	69 (13.8%)	5 (31.2%)	
Missing	20	19	1	
Hepatic encephalopathy[1]				0.44
No	277 (53.2%)	270 (53.5%)	7 (43.8%)	
Yes	244 (46.8%)	9 (56.2%)	235 (46.5%)	

Missing	16	15	1	
Previous hepatorenal syndrome[1]				0.52
No	499 (97.5%)	15 (100.0%)	484 (97.4%)	
Yes	13 (2.5%)	0 (0.0%)	13 (2.6%)	
Missing	25	23	2	
Antibiotic therapy for > 24 h[1]				0.52
No	486 (97.4%)	15 (100.0%)	471 (97.3%)	
Yes	13 (2.6%)	0 (0.0%)	13 (2.7%)	
Missing	38	36	2	
Hospitalized for > 48 h[1]				0.70
No	476 (95.4%)	462 (95.5%)	14 (93.3%)	
Yes	23 (4.6%)	22 (4.5%)	1 (6.7%)	
Missing	38	36	2	
Pre-transplant hemodialysis[1]				0.01
No	486 (97.0%)	473 (97.3%)	13 (86.7%)	
Yes	15 (3.0%)	13 (2.7%)	2 (13.3%)	
Missing	36	34	2	
Hepatocellular carcinoma[1]				0.34
No	330 (63.7%)	318 (63.3%)	12 (75.0%)	
Yes	188 (36.3%)	184 (36.7%)	4 (25.0%)	
Missing	19	18	1	
Blood group[1]				0.39
O	230 (42.9%)	221 (42.6%)	9 (52.9%)	
A	231 (43.1%)	226 (43.5%)	5 (29.4%)	
B	53 (9.9%)	50 (9.6%)	3 (17.6%)	
AB	22 (4.1%)	22 (4.2%)	0 (0.0%)	
Missing	1	1	0	
Congestive heart failure[1]				0.69
No	518 (99.0%)	502 (99.0%)	16 (100.0%)	
Yes	5 (1.0%)	5 (1.0%)	0 (0.0%)	
Missing	14	13	1	
Previous angioplasty[1]				0.75
No	520 (99.4%)	504 (99.4%)	16 (100.0%)	
Yes	3 (0.6%)	3 (0.6%)	0 (0.0%)	
Missing	14	13	1	
Dyslipidemia[1]				0.53
No	512 (97.7%)	496 (97.6%)	16 (100.0%)	
Yes	12 (2.3%)	12 (2.4%)	0 (0.0%)	
Missing	13	12	1	
Systemic arterial hypertension[1]				0.04
No	388 (73.6%)	379 (74.3%)	9 (52.9%)	
Yes	139 (26.4%)	131 (25.7%)	8 (47.1%)	
Missing	10	10	0	

Familiar history of coronary artery disease[1]				0.43
No	416 (79.8%)	402 (79.6%)	14 (87.5%)	
Yes	105 (20.2%)	103 (20.4%)	2 (12.5%)	
Missing	16	15	1	
Previous acute myocardial infarction[1]				0.06
No	515 (98.7%)	501 (98.8%)	14 (93.3%)	
Yes	7 (1.3%)	6 (1.2%)	1 (6.7%)	
Missing	15	13	2	
Previous stroke[1]				0.38
Ischemic	3 (0.6%)	3 (0.6%)	0 (0.0%)	
Hemorrhagic	27 (5.2%)	25 (4.9%)	2 (12.5%)	
No	493 (94.3%)	479 (94.5%)	14 (87.5%)	
Missing	14	13	1	
Diabetes mellitus[1]				0.21
No	396 (75.6%)	386 (76.0%)	10 (62.5%)	
Yes	128 (24.4%)	122 (24.0%)	6 (37.5%)	
Missing	13	12	1	
Valve replacement[1]				0.96
Metallic	1 (0.2%)	1 (0.2%)	0 (0.0%)	
Biological	2 (0.4%)	2 (0.5%)	0 (0.0%)	
No	445 (99.3%)	435 (99.3%)	10 (100.0%)	
Missing	89	82	7	
Mitral insufficiency[1]				0.001
Mild	137 (28.8%)	131 (28.4%)	6 (46.2%)	
Moderate	3 (0.6%)	2 (0.4%)	1 (7.7%)	
Absent	335 (70.5%)	329 (71.2%)	6 (46.2%)	
Missing	62	58	4	
Tricuspid insufficiency[1]				0.88
Mild	23 (4.9%)	22 (4.8%)	1 (7.7%)	
Moderate	1 (0.2%)	1 (0.2%)	0 (0.0%)	
Absent	448 (94.9%)	436 (95.0%)	12 (92.3%)	
Missing	65	61	4	
Noninvasive diagnostic method for myocardial ischemia[1]				0.001
Negative	65 (12.8%)	58 (11.8%)	7 (41.2%)	
Positive	442 (87.2%)	432 (88.2%)	10 (58.8%)	
Missing	30	30	0	
Dynamic myocardial perfusion scintigraphy induced ischemia[1]				0.01
Negative	304 (88.9%)	292 (89.8%)	12 (70.6%)	
Positive	38 (11.1%)	33 (10.2%)	5 (29.4%)	
Missing	195	195	0	
Weight (kg)[2]				0.74
Mean (SD)	74.2 (14.2)	74.2 (14.4)	73.1 (10.2)	
Missing	9	9	0	

Height (cm)[2]				0.38
Mean (SD)	168.0 (9.0)	168.1 (9.1)	166.2 (6.9)	
Missing	12	12	0	
Body mass index (kg/m²)[2]				0.8
Mean (SD)	26.2 (4.3)	26.2 (4.3)	26.5 (3.5)	
Missing	13	13	0	
Hematocrit (%)[2]				0.98
Mean (SD)	36.2 (5.8)	36.2 (5.8)	36.2 (3.8)	
Missing	3	3	0	
White blood cell count (mm³)[2]				0.36
Mean (SD)	5153 (4539)	5185 (4602)	4163 (1473)	
Missing	2	2	0	
Platelets[2]				0.67
Mean (SD)	84524 (55946)	84340 (56004)	90058 (55570)	
Missing	9	0	0	
Total bilirubin (mg/dL)[2]				0.09
Mean (SD)	2.9 (3.4)	2.9 (3.5)	1.5 (0.7)	
Missing	1	0	0	
Creatinine (mg/dL)[2]				0.08
Mean (SD)	1.2 (1.8)	1.2 (1.7)	2.0 (2.8)	
Missing	5	5	0	
International normalized ratio[2]				0.50
Mean (SD)	1.5 (0.4)	1.5 (0.4)	1.5 (0.3)	
Missing	32	32	0	
Sodium (mEq/L)[2]				0.91
Mean (SD)	137.7 (4.9)	137.7 (4.9)	137.8 (4.6)	
Missing	37	36	1	
Potassium (mmol/L)[2]				0.90
Mean (SD)	4.4 (0.6)	4.4 (0.6)	4.4 (0.7)	
Missing	44	43	1	
Albumin (g/dL)[2]				0.64
Mean (SD)	3.2 (0.6)	3.2 (0.6)	3.3 (0.4)	
Missing	13	13	0	
Aspartate aminotransferase (U/L)[2]				0.15
Mean (SD)	89.9 (60.1)	90.6 (60.7)	69.0 (34.8)	
Missing	9	8	1	
Alanine aminotransferase (U/L)[2]				0.62
Mean (SD)	74.1 (62.8)	74.3 (62.9)	66.5 (60.7)	
Missing	8	7	1	
Gamma-glutamyl transferase (U/L)[2]				0.56
Mean (SD)	102.8 (131.8)	102.3 (131.5)	122.6 (145.0)	
Missing	33	30	3	
Alkaline phosphatase (U/L)[2]				0.75

Mean (SD)	147.2 (127.9)	146.9 (128.4)	157.6 (113.2)	
Missing	24	22	2	
Alpha-fetoprotein (ng/mL)[2]				0.67
Mean (SD)	49.6 (339.5)	50.8 (345.2)	15.2 (26.7)	
Missing	20	20	0	
Fasting blood glucose (mg/dL)[2]				0.65
Mean (SD)	108.6 (41.7)	108.7 (42.1)	104.1 (29.2)	
Missing	26	26	0	

[1]n (%), Pearson's χ^2 test.
[2]Mean (SD), linear model ANOVA.
MACE: Major adverse cardiovascular event.

Model explanations

Figure 3 presents feature importance analysis as per mean SHAP values. It reveals that, at the cohort-wide level, the most significant variables for prediction of postoperative MACE were negative noninvasive cardiac stress testing, use of a nonselective beta-blocker, direct bilirubin levels, blood type O, and dynamic alterations on MPS. SHAP values are averaged, and the impact of each feature on individual predictions may vary. For instance, the feature 'blood type O' may have varying impacts depending on the specific conditions and characteristics of the patient.

DISCUSSION

The aim of the present study was to assess the risk of in-hospital post-LT MACE and identify clinically relevant predictors of such events. In pursuit of this objective, we constructed a machine learning-based risk stratification model which could be made available online to assist clinicians in identifying LT recipients at heightened cardiac risk immediately after LT. These models hold significance due to cardiovascular causes being a leading contributor to post-LT mortality, and the absence of risk prediction models tailored to patients with ESLD.

In this study, various recipient-related factors known prior to LT were thoroughly examined. An optimized clinical model demonstrated predictive capabilities for in-hospital MACE following LT, exhibiting a strong discriminative performance with an area under the curve (AUC) of 0.89. This surpasses the performance reported in a previously published study attempting to predict similar outcomes, which achieved an AUC of 0.71[45].

The present study employed a comprehensive set of candidate variables gathered during the pre-LT evaluation, which encompassed a wide array of cardiovascular risk factors. Notably, the machine learning model consistently demonstrated superior performance across all endpoints, highlighting significant improvements when compared to widely utilized traditional models.

On performance analysis, the XGBoost model demonstrated remarkable predictive capability, achieving an impressive AUROC of 0.89. This performance highlights its effectiveness in predicting postoperative MACE in our cohort of 575 LT patients. Furthermore, our classification results revealed excellent precision (0.89), recall (0.80), and an F1-score of 0.84 for the negative class, underscoring the model's precision in identifying patients at low risk of MACE. The exceptional performance of the model is further substantiated by the calibration results, where the isotonic-calibrated model achieved optimal calibration, as indicated by the lowest Brier score of 0.100. This calibration ensures a high level of precision, recall, F1-score, and accuracy for both negative and positive classes, aligning the model's predictions closely with observed outcomes. The calibration curve (available as supplemental material) visually depicts the model's excellent calibration performance.

To gain insights into the factors influencing postoperative MACE in our cohort, we conducted feature importance analysis, as depicted in Figure 3. Our analysis revealed that several variables - namely, outcomes of noninvasive cardiac stress testing, administration of nonselective beta-blockers, direct bilirubin levels, blood type O, and dynamic alterations on MPS - contributed significantly to prediction of postoperative MACE at the cohort-wide level. These findings emphasize the importance of considering both cardiac and liver-related factors in assessing the risk of post-transplant MACE. It bears stressing that, while these variables hold substantial predictive power at the cohort level, their impact may vary for individual patients, depending on their unique clinical characteristics and conditions.

We also evaluated the performance of our models in comparison to existing cardiovascular disease risk prediction models, such as the Cardiovascular Risk in Orthotopic Liver Transplantation (CVROLT) score, which was derived from a cohort of 1024 first-time LT recipients[8]. The CVROLT score included a multitude of donor- and recipient-related factors and identified pre-transplant heart failure, atrial fibrillation, diabetes, and the presence of respiratory failure at the time of transplantation as the most significant predictors of post-LT adverse cardiovascular events. Notably, our study used similar source variables but employed advanced machine learning techniques, which, uniquely, allowed our models to be internally validated in a series of "blinded" test cohorts, enhancing the generalizability of the results. While the CVROLT score achieved a C statistic of 0.78, our models demonstrated substantial predictive capability, particularly the XGBoost
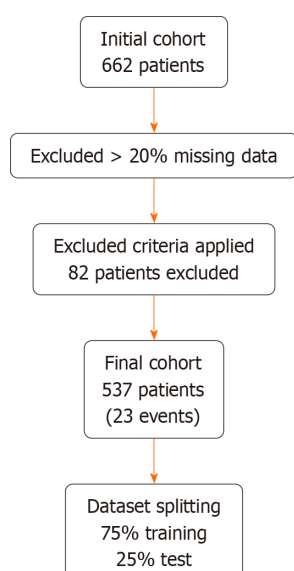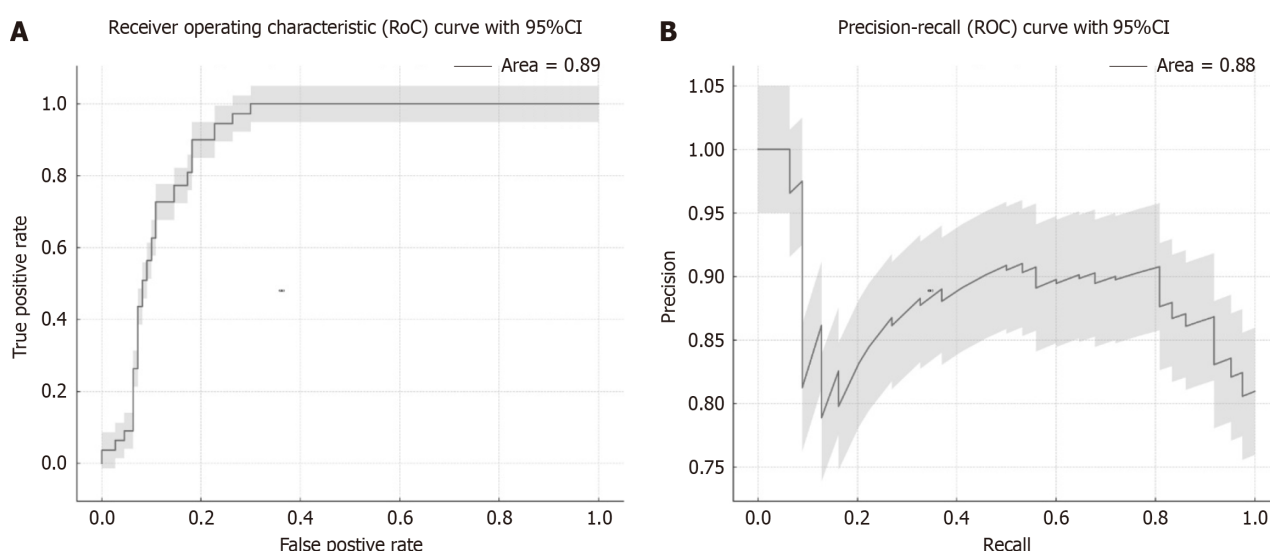
Initial cohort
662 patients

↓

Excluded > 20% missing data

↓

Excluded criteria applied
82 patients excluded

↓

Final cohort
537 patients
(23 events)

↓

Dataset splitting
75% training
25% test

Figure 1 Flow diagram of patient inclusion.

A

Receiver operating characteristic (RoC) curve with 95%CI

——— Area = 0.89

True positive rate (y-axis: 0.0, 0.2, 0.4, 0.6, 0.8, 1.0)

False positive rate (x-axis: 0.0, 0.2, 0.4, 0.6, 0.8, 1.0)

B

Precision-recall (ROC) curve with 95%CI

——— Area = 0.88

Precision (y-axis: 0.75, 0.80, 0.85, 0.90, 0.95, 1.00, 1.05)

Recall (x-axis: 0.0, 0.2, 0.4, 0.6, 0.8, 1.0)

Figure 2 Area under the receiver operating characteristic curve and area under the precision-recall curve for the model on the validation set. A and B: The area under the receiver operating characteristic curve in Figure 2A plots the true positive rate (sensitivity) against the false positive rate (1-specificity) for various threshold values. The area under the precision-recall curve in Figure 2B illustrates precision × recall for different threshold values. The shaded region represents the 95% confidence interval in both figures. ROC: Receiver operating characteristic; CI: Confidence interval.

model (AUC = 0.89). As noted above, this exceptional performance underscores the superiority of our models in predicting postoperative MACE in the context of LT.

The Revised Cardiac Risk Index (RCRI), another model traditionally used for predicting postoperative cardiovascular risk in individuals undergoing noncardiac surgery, has limited applicability in LT candidates[46]. The RCRI derivation cohort excluded patients with ESLD and primarily aimed to detect underlying ischemic heart disease, resulting in a suboptimal tool for risk-stratifying LT candidates for the occurrence of long-term MACE.

Both Josefsson *et al*[47] and Umphrey *et al*[48] reported on smaller cohorts of LT patients (*n* = 202 and *n* = 157, respectively). In their study, Josefsson *et al*[47] identified renal impairment, prolonged QTc, and age > 52 years as predictors of 1-year cardiovascular mortality. Similarly, Umphrey *et al*[48] investigated the role of DSE and reported that maximum heart rate achieved during the procedure, together with the model for end-stage liver disease (MELD) score, may predict adverse cardiovascular events up to 4 months post-orthotopic LT. Both of these previous models were limited by relatively small sample sizes, which may have impacted their external validity.

Historically, the assessment of cardiovascular risk in LT candidates has often prioritized the evaluation of CAD using methods such as DSE or coronary artery calcium scoring. This focus was largely driven by the high prevalence of traditional cardiovascular risk factors in LT recipients. However, the landscape is evolving as transplantation is increasingly performed on a medically complex population with higher median age at transplantation and higher MELD
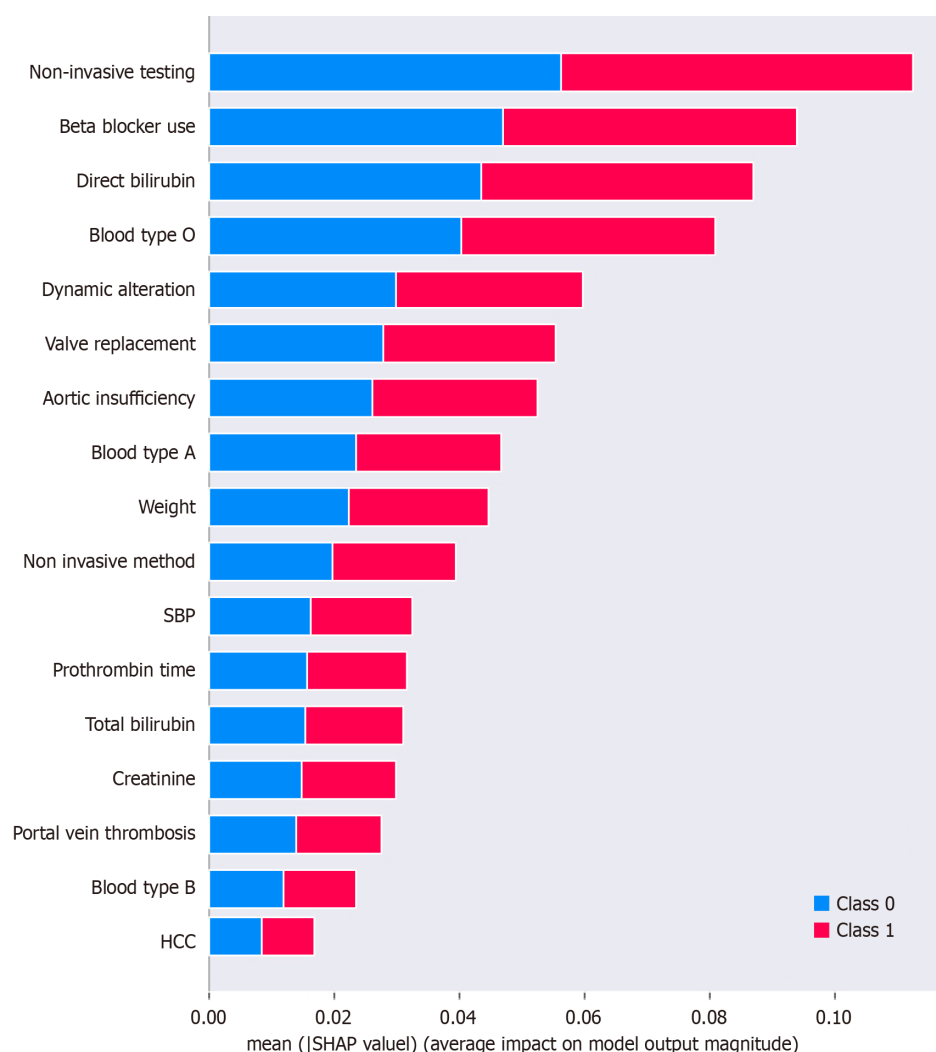
Figure 3 The x-axis represents the mean Shapley additive explanations value, which quantifies the average impact of each feature on the model's output. A higher mean Shapley additive explanations value means that the feature has a more significant influence on model predictions. The bars are color-coded to represent two distinct classes: Class 0 (blue), which represents absent major adverse cardiovascular event (MACE), and Class 1 (red), which represents the occurrence of MACE. The length of the bar in each color indicates the average impact of the corresponding feature on prediction of that specific class. Longer bars (regardless of color) mean that the feature has a greater average impact on model output. The direction of the influence (whether it pushes predictions towards Class 0 or Class 1) is denoted by the color. SBP: Spontaneous bacterial peritonitis; HCC: Hepatocellular carcinoma; Class 0: Major adverse cardiovascular event absent; Class 1: Major adverse cardiovascular event present; SHAP: Shapley additive explanations.

scores. Notably, advanced age alone correlates with cardiovascular comorbidities and independently predicts adverse cardiovascular events[1]. Additionally, ESLD is characterized by a high-output state with compromised ventricular reserve, known as cirrhotic cardiomyopathy, which may be exacerbated by the hemodynamic stress of liver reperfusion.

Recent systematic reviews and meta-analyses have shed light on the value of DSE in patients listed for LT. These studies reported that DSE had variable sensitivity (ranging from 20% to 32%) and specificity (ranging from 78% to 99%) for detecting CAD[25,26,49,50] mixed predictive capabilities for MACE post-LT, with sensitivity ranging from 20% to 28% and specificity as low as 78%[25,26,48,49]. It is evident that, while DSE exhibits a high negative predictive value, it may not be a reliable test for detecting risk of cardiovascular events, mortality, or presence of CAD in LT candidates. Therefore, its use should be reserved for selected intermediate-risk patients[51-53].

Furthermore, Oprea-Lager *et al*[54] demonstrated that the presence of a reversible perfusion defect suggestive of myocardial ischemia on MPS appears to increase all-cause mortality post-LT, with a hazard ratio of 3.17. Regarding MPS, several systematic reviews and meta-analyses have been conducted to evaluate its value in LT candidates. One such analysis, including five studies, found that MPS had a sensitivity of 62% and a specificity of 83% for detecting CAD[50]. Another diagnostic meta-analysis, involving 10 studies, reported a sensitivity of 82% and a specificity of 74% for MPS in CAD detection[25]. Finally, a prognostic meta-analysis revealed that positive MPS was associated with a relative risk of 2.6 (95%CI: 1.09-6.1) for major cardiac events and a relative risk of 2.7 (95%CI: 1.25-5.9) for mortality post-LT[26].

In patients listed for LT, the presence of coronary calcium has been significantly associated with various factors, including age, systolic blood pressure, alcohol-related cirrhosis, fasting blood glucose levels, the number of metabolic syndrome criteria, and the number of affected vessels. Importantly, coronary artery calcium score (CACS) values offer valuable insights into cardiac risk stratification. A CACS below 100 predicts a very low risk of post-LT cardiac events,

while a CACS above 250 suggests the need for coronary angiography[55] and a CACS exceeding 400 identifies patients at risk of MACE for up to 5 years post-LT. A recent study from 2021, comparing the diagnostic accuracy of DSE and CACS in detecting CAD, demonstrated the superiority of CACS over DSE[56].

Currently, it is proposed that coronary computed tomography angiography (CCTA) serves as the initial testing strategy for LT candidates with moderate to high CAD risk, while low-risk patients may not require additional cardiovascular assessment[51]. However, it is essential to acknowledge that CCTA may have limitations in detecting functional microvascular disease, which can contribute to type 2 myocardial infarction post-LT[57].

A recent systematic review has highlighted the promising role of machine learning models in improving prognostication for LT. The authors have found that machine learning models consistently outperformed traditional scoring systems, demonstrating excellent predictive capabilities for various post-transplant complications, including mortality, sepsis, and acute kidney injury. They suggest that machine learning could enhance decision-making related to organ allocation and LT, representing a substantial advancement in prognostication methods[58].

In the future, generalist medical artificial intelligence (GMAI) may bring a paradigm shift in medical AI use. Emphasizing flexibility and reusability, GMAI models can perform diverse tasks with minimal labeled data, developed through self-supervision on extensive datasets[59]. This might cause a shift in this paradigm, driven by hardware advances and the demand for personalized care, emphasizing AI's role in decision-making and improving diagnostic and prognostic performance[60].

In the context of utilizing machine learning to predict major MACE following LT, addressing the ethical implications and challenges that arise when implementing these models in clinical practice is crucial. The integration of machine learning introduces concerns surrounding data privacy, as patient information must be handled securely to protect confidentiality. Additionally, ensuring model transparency is essential, as clinicians need to understand the decision-making process of the machine learning model to trust its predictions. Furthermore, the potential biases embedded in the training data used for these models must be carefully examined and mitigated to avoid disproportionate effects on certain patient populations. By discussing these ethical considerations, the application of machine learning in predicting post-LT MACE can be approached with a well-rounded perspective that prioritizes patient privacy, model transparency, and fairness in healthcare outcomes.

This study is subject to several limitations. The retrospective design introduces inherent biases and data limitations. Significantly, a notable portion of the excluded patients, marked by a substantial volume of missing data, underwent LT with increased celerity attributed to higher MELD scores, and this resulted in an incomplete pre-LT clinical or cardiological evaluation. Second, the single-center setting may limit the generalizability of the findings to broader patient populations. Third, it is important to note that, while the machine learning model provides valuable predictive insights, it should serve as an aid to clinical judgment rather than a replacement, as it is better suited to predict a general rather than an individual risk of MACE. Additionally, the exclusion of certain patient groups based on specific criteria may impact the model's applicability in real-world scenarios. Finally, while the SHAP framework offers insights into feature importance, further investigation is needed to establish clinical relevance. While the study presents a robust predictive model, these limitations should be taken into consideration when interpreting and applying its results; future research with a view to external validation and improvement of clinical utility will be welcome.

The uncertainty surrounding the positive or negative outcomes of noninvasive tests and the prevalence of blood type O as risk factors for MACE highlights a critical aspect of machine learning model interpretability - it is advisable to avoid overestimating the significance and generalization of such information. The limitation of many models, including XGBoost, is the absence of clarity on why negative noninvasive cardiac stress testing correlates with a reduced risk of MACE. While these models excel at identifying statistical patterns, they often fall short in providing explicit explanations for correlations, lacking inherent insights into the biological or clinical reasons behind observed associations. Complementary research to unravel the biological significance of these correlations is required, emphasizing the distinction between mathematical patterns and causal relationships.

In this context, we can only speculate about these variables. Blood type O has shown a negative association with myocardial infarction[61-63], adding an intriguing dimension to the findings of the machine learning model. In patients with ESLD, distinguishing whether chronotropic incompetence results from cirrhosis-related autonomic dysfunction or is solely due to a beta-blocker effect is challenging. This ambiguity leads to numerous false negatives in stress testing, potentially influencing the negative association observed between stress testing and MACE[64]. One particularly intriguing discovery was the correlation between liver function markers and MACE - arguably the most noteworthy among these variables. Often, liver function is underestimated, and its impact on MACE may be overlooked, with attention primarily directed at the heart. Emphasizing the evaluation of both cardiac and hepatic aspects is crucial in pre-LT cardiac assessments[65].

The meticulous evaluation of pre-LT factors, incorporation of advanced machine learning techniques, and the demonstrated superior performance of the XGBoost model in predicting MACE distinguish this study. The model developed outperforms existing risk prediction tools, such as the CVROLT and CAR-OLT scores, and adds significant value to the relevant and current discussion on this topic. Additionally, the insights from this research not only contribute to the current knowledge but also pave the way for more accurate and tailored risk predictions in the context of LT.

CONCLUSION

In conclusion, the outcomes produced by our developed machine learning model are consistent with findings reported in prior literature. The calibration analysis indicates that our efforts to prevent overfitting and data leakage have indeed

been successful, suggesting that results are likely to remain stable when the model is applied to prospective data. Moreover, we have integrated the model into a user-friendly MACE prediction calculator which is now available online. This implementation will enable us to conduct a more comprehensive assessment of its prospective impact on prognosis.

With the increasing volume of LT procedures, the machine learning model presented herein can serve as a valuable resource for patient counseling, shared clinical decision-making with patient consent, quality improvement, and development of risk-reduction strategies. Further validation and application of this machine learning model in other registries and patient populations are essential to better understand its external validity in patients undergoing LT across multiple major transplantation-capable tertiary referral centers.

ARTICLE HIGHLIGHTS

Research background

The landscape of liver transplant (LT) candidates has evolved, with an aging and increasingly morbid population, often linked to metabolic-associated fatty liver disease (MAFLD). MAFLD's rise as a cause of cirrhosis raises concerns about a subsequent increase in major adverse cardiovascular events (MACE) post-LT, a critical complication negatively impacting prognosis. This study is prompted by the growing incidence of post-LT MACE, particularly within the first 6 months, and the complex interplay of traditional and nontraditional cardiovascular risk factors in this vulnerable population. The prevalence shift toward MAFLD as a leading indication for LT necessitates a thorough pre-LT cardiac assessment, demanding a reconsideration of existing noninvasive strategies' reliability. The pressing need for an alternative approach to predict post-LT MACE accurately propels the exploration of machine learning as a transformative tool to navigate the challenges posed by conventional models.

Research motivation

Motivating this research is the imperative to address the limitations of current cardiovascular risk stratification models for LT candidates, especially those with end-stage liver disease. Traditional models exhibit constraints related to assumptions of linear relationships and limited variables, leading to unreliable predictions. The inadequacy of existing noninvasive strategies and the absence of effective models for accurate cardiovascular risk stratification in LT candidates underscore the urgency for a paradigm shift. The study is driven by the aspiration to introduce machine learning as an innovative and more effective approach, leveraging its capacity to discern intricate patterns and relationships within datasets. The ultimate goal is to revolutionize risk prediction, enabling clinicians to identify high-risk individuals with precision, thus optimizing patient care strategies.

Research objectives

The primary objective of this study is to assess the feasibility and accuracy of implementing a machine learning model to predict MACE post-LT. Focusing on a specific regional cohort, the study aims to revolutionize risk assessment by moving beyond the limitations of conventional statistical models. Realizing this objective involves scrutinizing the potential of machine learning techniques to forecast post-LT MACE with enhanced precision. By leveraging advanced computational models, the research seeks to provide a comprehensive evaluation of the predictive capabilities, enabling the early identification of individuals at elevated risk. The ultimate significance lies in facilitating early intervention strategies and refining patient care in the context of the evolving landscape of LT candidates.

Research methods

This retrospective cohort study, approved by the Research Ethics Committee, delves into the cardiovascular risks following LT. Employing a comprehensive approach, medical records from Irmandade Santa Casa de Misericórdia de Porto Alegre were scrutinized for patients undergoing their first LT between 2001 and 2011 due to cirrhosis. Rigorous inclusion and exclusion criteria were applied, focusing on patients above 18 years of age with complete records, cardiac evaluation pre-LT, and no retransplantation. Data encompassed pre-LT, perioperative, and post-LT periods, with the primary outcome being in-hospital MACE. Statistical analyses, including frequency, means, standard deviation, Pearson's χ^2 test, and linear model analysis of variance, were executed using R software. The study introduces a machine learning paradigm, leveraging the XGBoost model, known for handling imbalanced datasets. Feature engineering involved a two-step imputation process, incorporating patient demographics, medical history, and cardiac evaluations. Model training incorporated regularization and early-stop techniques, aiming to prevent overfitting. Hyperparameter optimization using the Optuna package and performance evaluation metrics, including area under the receiver operating characteristic curve (AUROC) and area under the precision-recall curve, ensured robustness. Calibration, model explanation through Shapley additive explanations values, and adherence to the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis statement further enriched the methodological rigor, ultimately culminating in web deployment and code availability for transparency and accessibility.

Research results

The study involved 662 LT patients, with 82 exclusions based on specific criteria. The final dataset included 537 samples, with 23 in-hospital MACE cases. The XGBoost model demonstrated substantial predictive capability, achieving an AUROC of 0.89. Precision, recall, and F1-score for the negative class were 0.89, 0.80, and 0.84, respectively. The overall incidence of MACE was 4.46%, with observed rates for stroke, new-onset heart failure, severe arrhythmia, and

myocardial infarction. The model achieved optimal calibration using the isotonic method with a Brier score of 0.100. Feature importance analysis revealed key predictors, including negative noninvasive cardiac stress testing, use of a nonselective beta-blocker, direct bilirubin levels, blood type O, and dynamic alterations on myocardial perfusion scintigraphy. The findings contribute a valuable machine learning model for predicting post-LT MACE, offering insights into specific risk factors and enhancing precision in identifying at-risk individuals. Remaining challenges involve addressing potential variability in feature impact across patients and further validation in diverse cohorts.

Research conclusions

This study pioneers a novel approach in assessing in-hospital post-LT MACE. The research introduces a machine learning-based risk stratification model, surpassing the predictive performance of existing models, particularly demonstrating an impressive area under the curve of 0.89 using the XGBoost model. The optimized clinical model considers recipient-related factors and provides valuable insights into predicting MACE, crucial for addressing the leading cause of post-LT mortality. The use of machine learning techniques, specifically XGBoost, brings substantial improvements over traditional models, enhancing risk stratification accuracy. This study highlights the importance of comprehensive pre-LT evaluation, considering a wide array of cardiovascular risk factors.

Research perspectives

Future research should focus on refining and expanding the machine learning model's application, considering external validation in diverse patient populations and healthcare settings. Addressing ethical implications and ensuring transparency in model application are imperative for integrating machine learning predictions into clinical practice. The study suggests the need for continued exploration into the biological significance of identified predictors, such as the intriguing correlation between blood type O and reduced MACE risk. The model's implementation in a user-friendly MACE prediction calculator opens avenues for prospective impact assessment, counseling, shared decision-making, and risk reduction strategies in the growing landscape of LT procedures. External validation and application in various transplantation-capable centers will enhance understanding of the model's broader utility.

FOOTNOTES

REFERENCES

1   **Xia VW**, Taniguchi M, Steadman RH. The changing face of patients presenting for liver transplantation. *Curr Opin Organ Transplant* 2008; **13**: 280-284 [PMID: 18685318 DOI: 10.1097/MOT.0b013e328300a070]

2   **Onzi G**, Moretti F, Balbinot SS, Balbinot RA, Soldera J. Hepatocellular carcinoma in non-alcoholic fatty liver disease with and without cirrhosis. *Hepatoma Res* 2019; **5**: 7 [DOI: 10.20517/2394-5079.2018.114]

3   **Gill MG**, Majumdar A. Metabolic associated fatty liver disease: Addressing a new era in liver transplantation. *World J Hepatol* 2020; **12**: 1168-1181 [PMID: 33442446 DOI: 10.4254/wjh.v12.i12.1168]

4   **Burra P**, Becchetti C, Germani G. NAFLD and liver transplantation: Disease burden, current management and future challenges. *JHEP Rep* 2020; **2**: 100192 [PMID: 33163950 DOI: 10.1016/j.jhepr.2020.100192]

5   **Younossi ZM**, Marchesini G, Pinto-Cortez H, Petta S. Epidemiology of Nonalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis: Implications for Liver Transplantation. *Transplantation* 2019; **103**: 22-27 [PMID: 30335697 DOI: 10.1097/TP.0000000000002484]

6   **VanWagner LB**, Lapin B, Levitsky J, Wilkins JT, Abecassis MM, Skaro AI, Lloyd-Jones DM. High early cardiovascular mortality after liver transplantation. *Liver Transpl* 2014; **20**: 1306-1316 [PMID: 25044256 DOI: 10.1002/lt.23950]

7   **VanWagner LB**, Serper M, Kang R, Levitsky J, Hohmann S, Abecassis M, Skaro A, Lloyd-Jones DM. Factors Associated With Major Adverse Cardiovascular Events After Liver Transplantation Among a National Sample. *Am J Transplant* 2016; **16**: 2684-2694 [PMID: 26946333 DOI: 10.1111/ajt.13779]

8   **Vanwagner LB**, Bhave M, Te HS, Feinglass J, Alvarez L, Rinella ME. Patients transplanted for nonalcoholic steatohepatitis are at increased risk for postoperative cardiovascular events. *Hepatology* 2012; **56**: 1741-1750 [PMID: 22611040 DOI: 10.1002/hep.25855]

9   **Koshy AN**, Gow PJ, Han HC, Teh AW, Jones R, Testro A, Lim HS, McCaughan G, Jeffrey GP, Crawford M, Macdonald G, Fawcett J, Wigg A, Chen JWC, Gane EJ, Munn SR, Clark DJ, Yudi MB, Farouque O. Cardiovascular mortality following liver transplantation: predictors and temporal trends over 30 years. *Eur Heart J Qual Care Clin Outcomes* 2020; **6**: 243-253 [PMID: 32011663 DOI: 10.1093/ehjqcco/qcaa009]

10 **Konerman MA**, Fritze D, Weinberg RL, Sonnenday CJ, Sharma P. Incidence of and Risk Assessment for Adverse Cardiovascular Outcomes After Liver Transplantation: A Systematic Review. *Transplantation* 2017; **101**: 1645-1657 [PMID: 28296809 DOI: 10.1097/TP.0000000000001710]

11 **Watt KD**, Pedersen RA, Kremers WK, Heimbach JK, Charlton MR. Evolution of causes and risk factors for mortality post-liver transplant: results of the NIDDK long-term follow-up study. *Am J Transplant* 2010; **10**: 1420-1427 [PMID: 20486907 DOI: 10.1111/j.1600-6143.2010.03126.x]

12 **Raval Z**, Harinstein ME, Skaro AI, Erdogan A, DeWolf AM, Shah SJ, Fix OK, Kay N, Abecassis MI, Gheorghiade M, Flaherty JD. Cardiovascular risk assessment of the liver transplant candidate. *J Am Coll Cardiol* 2011; **58**: 223-231 [PMID: 21737011 DOI: 10.1016/j.jacc.2011.03.026]

13 **Kaur H**, Premkumar M. Diagnosis and Management of Cirrhotic Cardiomyopathy. *J Clin Exp Hepatol* 2022; **12**: 186-199 [PMID: 35068798 DOI: 10.1016/j.jceh.2021.08.016]

14 **Izzy M**, Oh J, Watt KD. Cirrhotic Cardiomyopathy After Transplantation: Neither the Transient Nor Innocent Bystander. *Hepatology* 2018; **68**: 2008-2015 [PMID: 29672903 DOI: 10.1002/hep.30040]

15 **Alqahtani SA**, Fouad TR, Lee SS. Cirrhotic cardiomyopathy. *Semin Liver Dis* 2008; **28**: 59-69 [PMID: 18293277 DOI: 10.1055/s-2008-1040321]

16 **Madhwal S**, Atreja A, Albeldawi M, Lopez R, Post A, Costa MA. Is liver transplantation a risk factor for cardiovascular disease? A meta-analysis of observational studies. *Liver Transpl* 2012; **18**: 1140-1146 [PMID: 22821899 DOI: 10.1002/lt.23508]

17 **Albeldawi M**, Aggarwal A, Madhwal S, Cywinski J, Lopez R, Eghtesad B, Zein NN. Cumulative risk of cardiovascular events after orthotopic liver transplantation. *Liver Transpl* 2012; **18**: 370-375 [PMID: 22140067 DOI: 10.1002/lt.22468]

18 **Fouad TR**, Abdel-Razek WM, Burak KW, Bain VG, Lee SS. Prediction of cardiac complications after liver transplantation. *Transplantation* 2009; **87**: 763-770 [PMID: 19295324 DOI: 10.1097/TP.0b013e318198d734]

19 **Eslam M**, Newsome PN, Sarin SK, Anstee QM, Targher G, Romero-Gomez M, Zelber-Sagi S, Wai-Sun Wong V, Dufour JF, Schattenberg JM, Kawaguchi T, Arrese M, Valenti L, Shiha G, Tiribelli C, Yki-Järvinen H, Fan JG, Grønbæk H, Yilmaz Y, Cortez-Pinto H, Oliveira CP, Bedossa P, Adams LA, Zheng MH, Fouad Y, Chan WK, Mendez-Sanchez N, Ahn SH, Castera L, Bugianesi E, Ratziu V, George J. A new definition for metabolic dysfunction-associated fatty liver disease: An international expert consensus statement. *J Hepatol* 2020; **73**: 202-209 [PMID: 32278004 DOI: 10.1016/j.jhep.2020.03.039]

20 **Ehtisham J**, Altieri M, Salamé E, Saloux E, Ollivier I, Hamon M. Coronary artery disease in orthotopic liver transplantation: pretransplant assessment and management. *Liver Transpl* 2010; **16**: 550-557 [PMID: 20440764 DOI: 10.1002/lt.22035]

21 **Targher G**, Byrne CD, Lonardo A, Zoppini G, Barbui C. Non-alcoholic fatty liver disease and risk of incident cardiovascular disease: A meta-analysis. *J Hepatol* 2016; **65**: 589-600 [PMID: 27212244 DOI: 10.1016/j.jhep.2016.05.013]

22 **Martin P**, DiMartini A, Feng S, Brown R Jr, Fallon M. Evaluation for liver transplantation in adults: 2013 practice guideline by the American Association for the Study of Liver Diseases and the American Society of Transplantation. *Hepatology* 2014; **59**: 1144-1165 [PMID: 24716201 DOI: 10.1002/hep.26972]

23 **Lentine KL**, Costa SP, Weir MR, Robb JF, Fleisher LA, Kasiske BL, Carithers RL, Ragosta M, Bolton K, Auerbach AD, Eagle KA; American Heart Association Council on the Kidney in Cardiovascular Disease and Council on Peripheral Vascular Disease; American Heart Association; American College of Cardiology Foundation. Cardiac disease evaluation and management among kidney and liver transplantation candidates: a scientific statement from the American Heart Association and the American College of Cardiology Foundation: endorsed by the American Society of Transplant Surgeons, American Society of Transplantation, and National Kidney Foundation. *Circulation* 2012; **126**: 617-663 [PMID: 22753303 DOI: 10.1161/CIR.0b013e31823eb07a]

24 **Ripoll C**, Yotti R, Bermejo J, Banares R. The heart in liver transplantation. *J Hepatol* 2011; **54**: 810-822

25 **Soldera J**, Camazzola F, Rodríguez S, Brandão A. Cardiac stress testing and coronary artery disease in liver transplantation candidates: Meta-analysis. *World J Hepatol* 2018; **10**: 877-886 [PMID: 30533188 DOI: 10.4254/wjh.v10.i11.877]

26   **Soldera J**, Camazzola F, Rodríguez S, Brandão A. Dobutamine stress echocardiography, myocardial perfusion scintigraphy, invasive coronary angiography, and post-liver transplantation events: Systematic review and meta-analysis. *Clin Transplant* 2018; **32**: e13222 [PMID: 29436036 DOI: 10.1111/ctr.13222]

27   **Kim WR**, Lake JR, Smith JM, Schladt DP, Skeans MA, Harper AM, Wainright JL, Snyder JJ, Israni AK, Kasiske BL. OPTN/SRTR 2016 Annual Data Report: Liver. *Am J Transplant* 2018; **18** Suppl 1: 172-253 [PMID: 29292603 DOI: 10.1111/ajt.14559]

28   **Safadi A**, Homsi M, Maskoun W, Lane KA, Singh I, Sawada SG, Mahenthiran J. Perioperative risk predictors of cardiac outcomes in patients undergoing liver transplantation surgery. *Circulation* 2009; **120**: 1189-1194 [PMID: 19752326 DOI: 10.1161/CIRCULATIONAHA.108.847178]

29   **Rachwan RJ**, Kutkut I, Timsina LR, Bou Chaaya RG, El-Am EA, Sabra M, Mshelbwala FS, Rahal MA, Lacerda MA, Kubal CA, Fridell JA, Ghabril MS, Bourdillon PD, Mangus RS. CAD-LT score effectively predicts risk of significant coronary artery disease in liver transplant candidates. *J Hepatol* 2021; **75**: 142-149 [PMID: 33476745 DOI: 10.1016/j.jhep.2021.01.008]

30   **VanWagner LB**, Ning H, Whitsett M, Levitsky J, Uttal S, Wilkins JT, Abecassis MM, Ladner DP, Skaro AI, Lloyd-Jones DM. A point-based prediction model for cardiovascular risk in orthotopic liver transplantation: The CAR-OLT score. *Hepatology* 2017; **66**: 1968-1979 [PMID: 28703300 DOI: 10.1002/hep.29329]

31   **Izzy M**, VanWagner LB, Lin G, Altieri M, Findlay JY, Oh JK, Watt KD, Lee SS; Cirrhotic Cardiomyopathy Consortium. Redefining Cirrhotic Cardiomyopathy for the Modern Era. *Hepatology* 2020; **71**: 334-345 [PMID: 31342529 DOI: 10.1002/hep.30875]

32   **Plotkin JS**, Scott VL, Pinna A, Dobsch BP, De Wolf AM, Kang Y. Morbidity and mortality in patients with coronary artery disease undergoing orthotopic liver transplantation. *Liver Transpl Surg* 1996; **2**: 426-430 [PMID: 9346688 DOI: 10.1002/Lt.500020604]

33   **Wray C**, Scovotti JC, Tobis J, Niemann CU, Planinsic R, Walia A, Findlay J, Wagener G, Cywinski JB, Markovic D, Hughes C, Humar A, Olmos A, Sierra R, Busuttil R, Steadman RH. Liver transplantation outcome in patients with angiographically proven coronary artery disease: a multi-institutional study. *Am J Transplant* 2013; **13**: 184-191 [PMID: 23126562 DOI: 10.1111/j.1600-6143.2012.04293.x]

34   **Deo RC**. Machine Learning in Medicine. *Circulation* 2015; **132**: 1920-1930 [PMID: 26572668 DOI: 10.1161/CIRCULATIONAHA.115.001593]

35   **Briganti G**, Le Moine O. Artificial Intelligence in Medicine: Today and Tomorrow. *Front Med (Lausanne)* 2020; **7**: 27 [PMID: 32118012 DOI: 10.3389/fmed.2020.00027]

36   **Soldera J**. Artificial intelligence as a prognostic tool for gastrointestinal tract pathologies. *Med Vozandes* 2023; **34**: 9-14 [DOI: 10.48018/rmv.v34.i1.e]

37   **Ballotin VR**, Bigarella LG, Soldera J. Deep learning applied to the imaging diagnosis of hepatocellular carcinoma. *Artif Intell Gastrointest Endosc* 2021; **2**: 127-135 [DOI: 10.37126/aige.v2.i4.127]

38   **Soldera J**, Tomé F, Corso LL, Machado Rech M, Ferrazza AD, Terres AZ, Cini BT, Eberhardt LZ, Balensiefer JIL, Balbinot RS, Muscope ALF, Longen ML, Schena B, Rost GL Jr. , Furlan RG, Balbinot RA, Balbinot SS. Use of a Machine Learning Algorithm to Predict Rebleeding and Mortality for Oesophageal Variceal Bleeding in Cirrhotic Patients. *EMJ Gastroenterol* 2020; **9**: 46-48

39   **Rech MM**, Corso LL, Dal Bó EF, Tomé F, Terres AZ, Balbinot RS, Muscope AL, Eberhardt LZ, Cini BT, Balensiefer JI, Schena B, Longen ML, Rost GL Jr, Balbinot SS, Balbinot RA, Soldera J. Prospective Validation of a Neural Network Model for the Prediction of 1-Year Mortality in Cirrhotic Patients with Acute Esophageal Variceal Bleeding. *Gastroenterology* 2023 [DOI: 10.1016/S0016-5085(23)04120-3]

40   **Lee HC**, Yoon SB, Yang SM, Kim WH, Ryu HG, Jung CW, Suh KS, Lee KH. Prediction of Acute Kidney Injury after Liver Transplantation: Machine Learning Approaches vs. Logistic Regression Model. *J Clin Med* 2018; **7** [PMID: 30413107 DOI: 10.3390/jcm7110428]

41   **Soldera J**, Tomé F, Corso LL, Ballotin VR, Bigarella LG, Balbinot RS, Rodriguez S, Brandão AB, Hochhegger B. Predicting 30 and 365-Day Mortality After Liver Transplantation Using a Machine Learning Algorithm. *Gastroenterology* 2021; **160**: S-789 [DOI: 10.1016/S0016-5085(21)02602-0]

42   **Ambale-Venkatesh B**, Yang X, Wu CO, Liu K, Hundley WG, McClelland R, Gomes AS, Folsom AR, Shea S, Guallar E, Bluemke DA, Lima JAC. Cardiovascular Event Prediction by Machine Learning: The Multi-Ethnic Study of Atherosclerosis. *Circ Res* 2017; **121**: 1092-1101 [PMID: 28794054 DOI: 10.1161/CIRCRESAHA.117.311312]

43   **Kumar S**, Chong I. Correlation Analysis to Identify the Effective Data in Machine Learning: Prediction of Depressive Disorder and Emotion States. *Int J Environ Res Public Health* 2018; **15** [PMID: 30572595 DOI: 10.3390/ijerph15122907]

44   **Tekumalla VRR**.   When Silver Is As Good As Gold: Using Weak Supervision to Train Machine Learning Models on Social Media Data. Georgia State University, 2022

45   **Jain V**, Bansal A, Radakovich N, Sharma V, Khan MZ, Harris K, Bachour S, Kleb C, Cywinski J, Argalious M, Quintini C, Menon KVN, Nair R, Tong M, Kapadia S, Fares M. Machine Learning Models to Predict Major Adverse Cardiovascular Events After Orthotopic Liver Transplantation: A Cohort Study. *J Cardiothorac Vasc Anesth* 2021; **35**: 2063-2069 [PMID: 33750661 DOI: 10.1053/j.jvca.2021.02.006]

46   **Lee TH**, Marcantonio ER, Mangione CM, Thomas EJ, Polanczyk CA, Cook EF, Sugarbaker DJ, Donaldson MC, Poss R, Ho KK, Ludwig LE, Pedan A, Goldman L. Derivation and prospective validation of a simple index for prediction of cardiac risk of major noncardiac surgery. *Circulation* 1999; **100**: 1043-1049 [PMID: 10477528 DOI: 10.1161/01.cir.100.10.1043]

47   **Josefsson A**, Fu M, Björnsson E, Castedal M, Kalaitzakis E. Pre-transplant renal impairment predicts posttransplant cardiac events in patients with liver cirrhosis. *Transplantation* 2014; **98**: 107-114 [PMID: 24621533 DOI: 10.1097/01.TP.0000442781.31885.a2]

48   **Umphrey LG**, Hurst RT, Eleid MF, Lee KS, Reuss CS, Hentz JG, Vargas HE, Appleton CP. Preoperative dobutamine stress echocardiographic findings and subsequent short-term adverse cardiac events after orthotopic liver transplantation. *Liver Transpl* 2008; **14**: 886-892 [PMID: 18508373 DOI: 10.1002/lt.21495]

49   **Nguyen P**, Plotkin J, Fishbein TM, Laurin JM, Satoskar R, Shetty K, Taylor AJ. Dobutamine stress echocardiography in patients undergoing orthotopic liver transplantation: a pooled analysis of accuracy, perioperative and long term cardiovascular prognosis. *Int J Cardiovasc Imaging* 2013; **29**: 1741-1748 [PMID: 23974907 DOI: 10.1007/s10554-013-0275-x]

50   **Tiwari N**, Margapuri J, Katamreddy A, Jubbal S, Madan N. Diagnostic accuracy of cardiac testing for coronary artery disease in potential liver transplant recipients: A systematic review and meta-analysis. *Int J Cardiol Heart Vasc* 2021; **32**: 100714 [PMID: 33521238 DOI: 10.1016/j.ijcha.2021.100714]

51   **Robertson M**, Chung W, Liu D, Seagar R, O'Halloran T, Koshy AN, Horrigan M, Farouque O, Gow P, Angus P. Cardiac Risk Stratification in Liver Transplantation: Results of a Tiered Assessment Protocol Based on Traditional Cardiovascular Risk Factors. *Liver Transpl* 2021; **27**: 1007-1018 [PMID: 33606328 DOI: 10.1002/lt.26025]

52   **De Gasperi A**, Spagnolin G, Ornaghi M, Petrò L, Biancofiore G. Preoperative cardiac assessment in liver transplant candidates. *Best Pract Res Clin Anaesthesiol* 2020; **34**: 51-68 [PMID: 32334787 DOI: 10.1016/j.bpa.2020.02.002]

53   **Cheng XS**, VanWagner LB, Costa SP, Axelrod DA, Bangalore S, Norman SP, Herzog CA, Lentine KL; American Heart Association Council on the Kidney in Cardiovascular Disease and Council on Cardiovascular Radiology and Intervention. Emerging Evidence on Coronary Heart Disease Screening in Kidney and Liver Transplantation Candidates: A Scientific Statement From the American Heart Association: Endorsed by the American Society of Transplantation. *Circulation* 2022; **146**: e299-e324 [PMID: 36252095 DOI: 10.1161/CIR.0000000000001104]

54   **Oprea-Lager DE**, Sorgdrager BJ, Jukema JW, Scherptong RW, Ringers J, Coenraad MJ, van Hoek B, Stokkel MP. Clinical value of myocardial perfusion scintigraphy as a screening tool in liver transplant candidates. *Liver Transpl* 2011; **17**: 261-269 [PMID: 21384508 DOI: 10.1002/lt.22234]

55   **Choi JM**, Kong YG, Kang JW, Kim YK. Coronary Computed Tomography Angiography in Combination with Coronary Artery Calcium Scoring for the Preoperative Cardiac Evaluation of Liver Transplant Recipients. *Biomed Res Int* 2017; **2017**: 4081525 [PMID: 28164120 DOI: 10.1155/2017/4081525]

56   **Kleb C**, Jain V, Sheth C, Wolski K, Kapadia S, Grimm R, Desai M, Krishnaswamy A, Kassis N, Sheng C, Zheng H, Cywinski J, Menon KVN, Eghtesad B, Diago Uso T, Quintini C, Schoenhagen P, Harb SC, Sharma V, Fares M. Comparison of Coronary Artery Calcium Scoring with Dobutamine Stress Echo for Detection of Coronary Artery Disease Before Liver Transplantation. *Ann Transplant* 2021; **26**: e934163 [PMID: 34934037 DOI: 10.12659/AOT.934163]

57   **Pellikka PA**, Nagueh SF, Elhendy AA, Kuehl CA, Sawada SG; American Society of Echocardiography. American Society of Echocardiography recommendations for performance, interpretation, and application of stress echocardiography. *J Am Soc Echocardiogr* 2007; **20**: 1021-1041 [PMID: 17765820 DOI: 10.1016/j.echo.2007.07.003]

58   **Chongo G**, Soldera J. Use of machine learning models for the prognostication of liver transplantation: A systematic review. *World J Transplant* 2023; In press

59   **Moor M**, Banerjee O, Abad ZSH, Krumholz HM, Leskovec J, Topol EJ, Rajpurkar P. Foundation models for generalist medical artificial intelligence. *Nature* 2023; **616**: 259-265 [PMID: 37045921 DOI: 10.1038/s41586-023-05881-4]

60   **Souza Filho EM**, Fernandes FA, Soares CLA, Seixas FL, Santos AASMDD, Gismondi RA, Mesquita ET, Mesquita CT. Artificial Intelligence in Cardiology: Concepts, Tools and Challenges - "The Horse is the One Who Runs, You Must Be the Jockey". *Arq Bras Cardiol* 2020; **114**: 718-725 [PMID: 32491009 DOI: 10.36660/abc.20180431]

61   **Abegaz SB**. Human ABO Blood Groups and Their Associations with Different Diseases. *Biomed Res Int* 2021; **2021**: 6629060 [PMID: 33564677 DOI: 10.1155/2021/6629060]

62   **Dunne E**, Qi QM, Shaqfeh ES, O'Sullivan JM, Schoen I, Ricco AJ, O'Donnell JS, Kenny D. Blood group alters platelet binding kinetics to von Willebrand factor and consequently platelet function. *Blood* 2019; **133**: 1371-1377 [PMID: 30642918 DOI: 10.1182/blood-2018-06-855528]

63   **Timur AA**, Barnard J, Murugesan G, Gandhi S, Bhatt DL, Kottke-Marchant K. The relation between ABO blood types and clinical and platelet function parameters in patients who underwent percutaneous coronary intervention. *Coron Artery Dis* 2019; **30**: 51-58 [PMID: 30407210 DOI: 10.1097/MCA.0000000000000676]

64   **Rudzinski W**, Waller AH, Prasad A, Sood S, Gerula C, Samanta A, Koneru B, Klapholz M. New index for assessing the chronotropic response in patients with end-stage liver disease who are undergoing dobutamine stress echocardiography. *Liver Transpl* 2012; **18**: 355-360 [PMID: 22140006 DOI: 10.1002/lt.22476]

65   **Simon TG**, Kartoun U, Zheng H, Chan AT, Chung RT, Shaw S, Corey KE. MELD-Na score predicts incident major cardiovascular events, in patients with nonalcoholic fatty liver disease (NAFLD). *Hepatol Commun* 2017; **1**: 429-438 [PMID: 29085919 DOI: 10.1002/hep4.1051]