HHS Public Access

Author manuscript

Mayo Clin Proc Digit Health. Author manuscript; available in PMC 2024 July 11.

Published in final edited form as:

Mayo Clin Proc Digit Health. 2024 June ; 2(2): 221–230. doi:10.1016/j.mcpdig.2024.03.005.

Deep Learning—Based Prediction Modeling of Major Adverse Cardiovascular Events After Liver Transplantation

Ahmed Abdelhameed, PhD,
Harpreet Bhangu, MD,
Jingna Feng, MS,
Fang Li, PhD,
Xinyue Hu, MS,
Parag Patel, MD,
Liu Yang, MD,
Cui Tao

McWilliams School of Biomedical Informatics (A.A., J.F., F.L., X.H., C.T.), University of Texas Health Science Center at Houston, TX; and Department of Artificial Intelligence and Informatics (A.A., J.F., F.L., X.H., C.T.) and Department of Transplantation (H.B., P.P., L.Y.), Mayo Clinic, Jacksonville, FL.

Abstract

Objective: To validate deep learning models' ability to predict post-transplantation major adverse cardiovascular events (MACE) in patients undergoing liver transplantation (LT).

Patients and Methods: We used data from Optum's de-identified Clinformatics Data Mart Database to identify liver transplant recipients between January 2007 and March 2020. To predict post-transplantation MACE risk, we considered patients' demographics characteristics, diagnoses, medications, and procedural data recorded back to 3 years before the LT procedure date (index date). MACE is predicted using the bidirectional gated recurrent units (BiGRU) deep learning model in different prediction interval lengths up to 5 years after the index date. In total, 18,304 liver transplant recipients (mean age, 57.4 years [SD, 12.76]; 7158 [39.1%] women) were used to develop and test the deep learning model's performance against other baseline machine learning models. Models were optimized using 5-fold cross-validation on 80% of the cohort, and model performance was evaluated on the remaining 20% using the area under the receiver operating characteristic curve (AUC-ROC) and the area under the precision-recall curve (AUC-PR).

Correspondence: Cui Tao, PhD, Department of Artificial Intelligence and Informatics, Mayo Clinic, 4500 San Pablo Road, Jacksonville, FL 32224 (Tao.Cui@mayo.edu).

POTENTIAL COMPETING INTERESTS

The authors report no competing interests.

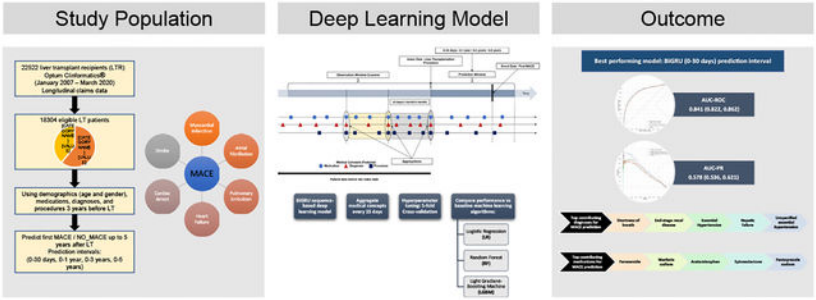
SUPPLEMENTAL ONLINE MATERIAL

Supplemental material can be found online at https://www.mcpdigitalhealth.org/. Supplemental material attached to journal articles has not been edited, and the authors take responsibility for the accuracy of all data.

Results: Using different prediction intervals after the index date, the top-performing model was the deep learning model, BiGRU, and achieved an AUC-ROC of 0.841 (95% CI, 0.822–0.862) and AUC-PR of 0.578 (95% CI, 0.537–0.621) for a 30-day prediction interval after LT.

Conclusion: Using longitudinal claims data, deep learning models can efficiently predict MACE after LT, assisting clinicians in identifying high-risk candidates for further risk stratification or other management strategies to improve transplant outcomes based on important features identified by the model.

GRAPHICAL ABSTRACT



Cardiovascular disease has become the primary cause of early mortality after liver transplantation (LT) in the United States, overtaking infections and graft failures.[1–8] Nowadays, liver transplant candidates tend to be older and present more cardiovascular comorbidities, mainly owing to the increasing prevalence of metabolic dysfunction-associated steatohepatitis (MASH) as the top indication for LT.[1,9] Accurate assessment of cardiovascular risks after LT is essential for optimal resource allocation and improved clinical outcomes. Current noninvasive tests, such as dobutamine stress echocardiography and nuclear medicine cardiac perfusion tests, demonstrate low sensitivity and specificity in diagnosing subclinical coronary and myocardial disease in cirrhotic patients.[3–5] Altered hemodynamics during LT can reveal hidden cardiovascular diseases either intraoperatively or immediately postoperatively.[4] Despite comprehensive preoperative evaluation, early cardiovascular death rates post-LT are 4 times higher than those in other high-risk noncardiac operations.[10]

Major adverse cardiovascular events (MACE), such as myocardial infarction, atrial fibrillation, pulmonary embolism, heart failure, cardiac arrest, and stroke, substantially contribute to morbidity and mortality after LT.[1,9,11] Several risk factors for unfavorable cardiovascular outcomes post-LT have been identified, including age, left ventricular hypertrophy, high blood pressure, dyslipidemia, tobacco use, a family history of coronary artery disease (CAD), MASH, and diabetes.[1–6,12] Risk stratification using these factors can help pinpoint high-risk patients for post-transplant cardiovascular complications, subjecting them to more invasive tests such as cardiac catheterization.[3,12] Moreover, identifying and risk-stratifying the patients susceptible to these life-threatening cardiovascular complications would help physicians plan preventive care and therapy, maximizing quality of life post-transplantation.[1,3] The wealth of patient information before transplantation provided by large-scale electronic health records, such as claim data, can greatly benefit the prediction of post-LT complications.

In this study, we aimed to develop and validate the predictive capabilities of machine learning algorithms, particularly deep learning models, for identifying individuals at risk of MACE after LT. This enables physicians to take preventive measures to reduce the likelihood of such complications occurring.

METHODS

Patients and Study Design

This study used Optum's deidentified Clinformatics Data Mart Database,[13] being derived from a large, adjudicated claim data warehouse. We conducted a retrospective analysis using insurance claims for 22,522 liver transplant recipients between January 2007 and March 2020. The Clinformatics Data Mart was queried for information including diagnoses, procedures, medications, and demographic characteristics, before and after the LT index date, defined as the day a patient underwent LT. The search criteria to retrieve the patient cohort can be found in Supplemental Table 1 (available online at https://www.mcpdigitalhealth.org/). In total, 1106 patients were excluded to ensure that all patients in this study were adults (18 years or older) at the time of the index date. Eight more patients were also excluded owing to missing gender information. Of the 21,416 patients deemed eligible for the study, 3112 patients were later removed owing to a lack of claim records before the LT index date. As a result, the final study cohort consisted of 18,304 patients (Figure 1A).

In the final patient cohort, we marked the date of the first diagnosis of any of the 6 MACE—heart failure, atrial fibrillation, stroke, pulmonary embolism, myocardial infarction, or cardiac arrest—for each patient, as the event date if occurred within 5 years after the index date. The definition of these major adverse cardiovascular complications using both the International Classification of Diseases, ninth revision (ICD-9) and tenth revision (ICD-10) codes can be found in Supplemental Table 2 (available online at https://www.mcpdigitalhealth.org/). The number of patients who experienced each of the 6 MACE in the cohort during these 5 years is depicted in the bar plot in Supplemental Figure 1 (available online at https://www.mcpdigitalhealth.org/). It should be noted that some patients experienced 1 or more of the 6 MACE after the LT on the same or different dates. The event date of interest in this study was the date of the first MACE.

Primary Study Outcome

The primary outcome of the study was the prediction of whether patients will develop any of the MACE within 4 different time intervals (30 days and 1, 3, and 5 years) after receiving an LT.

Deep Learning Models and Prediction Performance Evaluation

To leverage the power of deep learning sequence models and capture clinically relevant scenarios, we aggregated the occurrence of any medical concept over a predetermined time interval (eg, 15 days, 1 month, or 3 months) across a 3-year observation window before the LT index date as shown in the study design diagram (Figure 1B). This was done because insurance claims for these medical concepts can be filed at any time during the observation

window and are not always coexisting. The figure also shows the index date, which is marked as day 0. MACE was modeled in our study using several prediction window sizes, all starting at the index date. The prediction windows used to test our models in different scenarios are 0–30 days, 0–1 year, 0–3 years, and 0–5 years.

Model Architecture

The deep learning model used in this study was the bidirectional gated recurrent unit (BiGRU) model, a type of sequence-processing models.[14] BiGRUs are well suited for learning meaningful representations from sequence data during training because they consider the temporal dependence between the current input data at a particular time and its previous and subsequent counterparts. For our MACE prediction task, we compared BiGRU's performance against 3 traditional machine learning algorithms: logistic regression, random forest, and light gradient-boosting machine. Further information on the data preparation and implementation details for the BiGRU model and the machine learning models can be found in Supplemental Table 3 (available online at https://www.mcpdigitalhealth.org/).

Data Selection and Training

The dataset was randomly divided into a training set consisting of 80% of the patient cohort and a test set consisting of 20% of the patient cohort for each of the 4 prediction scenarios. Within each subset, the balance between the positive (MACE) and negative (NO_MACE) classes was maintained. This test set was held out for final validation and comparison of all models' performance. Hyperparameter tuning of the models was done using a stratified 5-fold cross-validation procedure on the training set. The best-performing model was then chosen and retrained on the entire training set before being tested on the test set. More details about the whole process are shown in Supplemental Figure 2 (available online at https://www.mcpdigitalhealth.org/).

Model Evaluation

The performance of the deep learning model on the test data set was compared with other machine learning models using the area under the receiver operating curve (AUC-ROC) and the area under the precision-recall curve (AUC-PR). In imbalanced datasets where the number of positive class samples is much lower than the number of negative class samples, AUC-PR can be the more useful performance measure because it is an indicator of how well the model can handle the positive samples. Bootstrapping was used to resample the test data set 500 times with replacement, 95% CIs were calculated, and the lowest and highest metrics values were reported.

Our deep learning model was also used to generate a list of ranked features by importance, with weights indicating how significantly each feature contributed to MACE prediction. This was achieved through the integrated gradient,[15] an interpretability or explainability technique for deep neural networks. Using this technique, we were able to provide a comprehensive ranking for all input features that contributed to the model prediction for the entire test set.

RESULTS

Patient Characteristics

The baseline characteristics of the patient cohort are summarized in Table 1. Of the patients with MACE, 37.7% were females (n=2436) and 62.3% males (n=4025); in contrast, patients without MACE comprised 39.9% females (n=4722) and 60.1% males (n=7121). Patients with MACE exhibited a higher prevalence of elevated body mass index (12.2%, n=791) than those without MACE (8.5%, n=1011), increased smoking rates (24.9% vs. 17.4%), and marked differences in diabetes, CAD, hypertension, myocardial infarction history, peripheral vascular disease, dyslipidemia, previous strokes, pulmonary hypertension, alcohol misuse, and MASH. It is worth noting that the number of patients with and without MACE varied across the 4 prediction windows of testing scenarios as shown in Supplemental Figure 3 (available online at https://www.mcpdigitalhealth.org/). In our study, we examined the incidence of MACE up to 5 years post-transplantation, whereas most studies focused on 30 days, 90 days, 4 months, and 1 year post-transplantation. Table 1 reports that 35.3% (n=6461) of liver transplant recipients experienced MACE, which was higher than the incidence reported in the studies we reviewed.

Prediction Performance of the Deep Learning Model Compared With Baseline Machine Learning Models

Table 2 displays the prediction performance metrics of the BiGRU deep learning model and 3 machine learning models for each of the 4 prediction scenarios. The input data for the BiGRU model were aggregated as a sequence of feature vectors every 15 days. The BiGRU model performed better than the 3 baseline models in all 4 testing scenarios except the 0- to 3-year scenario, where the light gradient-boosting machine model achieved a higher AUC-PR of 0.646 (0.62–0.673) than that of the BiGRU model (0.639 [0.613–0.661]).

Furthermore, we observed that the BiGRU model achieved the best AUC-ROC value in the 0- to 30-day testing scenario. This indicates that, when compared with longer prediction intervals, the model performed best in predicting MACE occurrence within the first 30 days post-LT. However, the AUC-PR values for the 0- to 1-year, 0- to 3-year, and 0- to 5-year testing scenarios, where the imbalance between the number of patients in MACE and NO_MACE groups decreased (Supplemental Figure 3), were found to be better than those for the 0- to 30-day scenario. Figure 2 shows the receiver operating characteristic and the precision-recall curves for the 4 models in the 0- to 30-day testing scenario. The curves emphasize that the deep learning model outperformed the baseline machine learning models for predicting MACE within the first 30 days after LT. Furthermore, the deep learning model obtained the best AUC-PR value (0.578), indicating that it was the best at addressing the imbalance between the number of patients with MACE and those without MACE in the test set.

To provide an interpretable model, rankings of the pre-transplant diagnosis and medication features for predicting MACE using the top-performing BiGRU model in the 0- to 30-day interval were calculated. Figure 3 shows the top 15 MACE predictors ranked by relevance in descending order for diagnoses and medications.

DISCUSSION

This study validated a deep learning model using pre-transplant data to predict MACE post-LT, analyzing over 20,000 patients with 6000þ MACE cases. It demonstrated the model's efficiency in forecasting short-term and long-term cardiovascular issues. Using extensive electronic health record data sets, including claims data, enhances post-transplant complication predictions by providing comprehensive pre-transplant patient information. The study sourced data from Optum's de-identified Clinformatics Data Mart Database, allowing broad analysis beyond single institutions. It highlighted the effectiveness of machine learning, especially deep learning, in handling large, complex datasets without extensive feature selection, using techniques like recurrent neural networks to explore temporal data and reveal critical clinical correlations.

Machine learning, especially neural networks, aids liver disease research, covering pre-transplant predictions (waitlist outcomes and hepatic steatosis), liver segmentation, graft allocation, and posttransplant forecasts (survival, rejection, failure, and postoperative risks).[16,17]

This pioneering model forecasts short-term and long-term MACE post-LT, using comprehensive pre-transplant data (age, gender, diagnoses, medications, and procedures). Leveraging patient claims, it proves deep learning's efficiency in predicting MACE and identifying key risk factors for transplant candidates.

Multiple studies have been published to provide predictive information for MACE after LT, with 4 papers using single-center data and 2 publications using large national sample databases.[1,6,18–20] However, the main limitation of single-center data is the small sample size, limited generalizability, and lack of calibration and status stick reporting.

Studies using large national data are limited by the assumption of linear correlation between pre-LT predictive variables and output variables, independence of predictive variables, dependence on a small number of selected predictors owing to concerns of over-fitting, and reliance on traditional statistical models such as logistic regression.[21–23]

Northwestern University's study introduced the CAR-OLT score to estimate cardiovascular risks post-LT, highlighting its effectiveness yet noting limitations.[6] CAR-OLT, confined to a year's post-LT risk assessment, contrasts with our deep learning model, extending predictions to 5 years. The model's reliance on the top 3 discharge codes for its risk equation might underrepresent cardiovascular complications. In addition, being limited to Northwestern Medicine hospitals could omit broader cardiovascular event data. The study's focus on a single center may also skew results toward specific practices and patient demographic characteristics. Furthermore, our investigation encompassed 2 alternative prediction models, broadening the scope of risk assessment tools in this field. Umphrey et al[19] created a model that used the percentage of maximum predicted heart rate and the rate pressure obtained during dobutamine stress echocardiographic along with model for end-stage liver disease score to predict cardiovascular events up to 4 months post-transplantation.[19] The study's main limitations are the small cohort of patients studied, the short posttransplant follow-up period, and the external validation required. Josefsson et al[18]

developed a score consisting of pre-transplant renal impairment, prolonged QTc interval, and older age, which would help identify the need for extensive cardiac testing in liver transplant candidates. The small sample size and the need for external validation are the study's main limitations.

We verified previously published high-risk factors associated with MACE, including age, gender, body mass index, smoking, diabetes, a history of CAD, hypertension, previous myocardial infarction, peripheral vascular disease, dyslipidemia, previous stroke, pulmonary hypertension, alcohol abuse, and MASH.[1,6,18–20] Our study stands out by examining a comprehensive set of clinical risk factors not collectively analyzed in previous research, with a notably larger patient cohort.[18,19]

We explored the BiGRU deep learning model and 3 machine learning models, assessing different observation windows (1 and 2 years) and aggregation intervals (1 and 3 months). Narrower observation windows with larger aggregation intervals degraded performance. Unlike previous models built on selected variables for cardiovascular risk or MACE prediction post-LT, we used all available pre-transplant data, avoiding selective bias. For instance, a study used 35 variables to assess machine learning models for MACE prediction in liver transplant patients, where XGBoost reported the best AUC-ROC of 0.71 (0.63–0.79).[23] Another study used up to 190 variables for deep learning models to predict fatal post-LT complications, achieving an AUC-ROC of 0.807 (0.795–0.842) for 1-year and 0.722 (0.705–0.764) for 5-year predictions.[21] Our approach leverages the full spectrum of pretransplant data, maximizing deep learning's potential with high-dimensional datasets.

To interpret our deep learning model's findings, we pinpointed key MACE predictors post-LT (Figure 3). Using a relative contribution score, we assessed the importance of each predictor within the model. The process begins by determining each predictor's absolute importance through its cumulative impact across the testing set. This importance is then normalized to a relative score, capped at 1, through minimum-maximum normalization. Subsequently, predictors related to MACE diagnoses and medications are ranked by their contribution. It is important to note that these scores indicate the strength of the predictor's association with MACE but not the direction (positive or negative) of the association. Key pre-transplant factors linked to MACE within the first 30 days post-LT include shortness of breath, end-stage renal disease, essential hypertension, and hepatic failure without coma, along with medications such as furosemide, warfarin, acetaminophen, spironolactone, and pantoprazole. Warfarin use, often related to atrial fibrillation or thrombosis history, and the prescription of diuretics for fluid management in cirrhosis or heart failure are noteworthy. The connection of acetaminophen and pantoprazole with MACE requires further exploration. These insights underscore the importance of monitoring specific conditions and medications to effectively manage MACE risks in the critical initial post-LT phase.

The strengths of this study include the use of a large patient population and a machine learning model that uses all available patient data as predictors for MACE without the need for expert feature selection. Our model can predict not only short-term but also long-term cardiovascular complications whereas not being restricted to one hospital or member

hospitals. Deep learning can identify complex relationships between high-dimensional data, which can help model complex multivariate functions. Furthermore, sequence-based deep learning models can leverage longitudinal data to learn temporal dependencies between the input variables (features) which can result in improved performance. To compensate for the lack of external validation, we established 2 levels of internal validation. The first level used 5-fold cross-validation using 80% of the cohort. At the second level, the model was validated using a completely independent test set comprising 20% of the cohort. Finally, using pre-transplantation longitudinal data, we created an interpretable model and were able to explore the most important input variables (in both categories, diagnoses, and medications) for predicting MACE after LT.

However, it is important to acknowledge the limitations of our study. First, we encountered common problems related to the quality of the dataset, including the presence of incorrect or missing data such as patient phenotypes and LT-related information such as ischemia/reperfusion time, operation time, quality of donor graft, and severity of liver disease (eg, model for end-stage liver disease [MELD] score). These issues may have introduced some degree of noise or bias into our analysis. We also acknowledge the intricacies posed by conflicting risks, such as organ rejection or recurrence of liver disease, which can impact the interpretation of MACE occurrences post-LT. In addition, our study is retrospective in nature and focused exclusively on pre-transplantation variables to predict post-transplantation MACE. Given the crucial nature of the operation and the chance of MACE occurring on the same day, we decided to align our prediction intervals with the day of LT. However, this approach restricts the consideration of post-transplant variables that could potentially influence MACE outcomes. Although we recognize the importance of testing our model's generalizability on a different dataset, our model has not been externally validated owing to a lack of accessibility. We attempted to address this limitation by implementing 2 levels of internal validation in our study design. Nevertheless, since our model uses standard coded pre-transplant variables, similar datasets could be easily adapted for validation without the need for retraining the model.

CONCLUSION

This study presented compelling evidence supporting the use of deep learning models as a risk-stratifying tool for predicting MACE in transplant recipients using pre-transplantation longitudinal claim data. Our findings indicate that leveraging all available patient data as predictors for MACE eliminates the need for expert feature selection to prespecify a certain set of features. This approach has the potential to aid clinicians in identifying high-risk transplant recipients and developing targeted interventions to reduce MACE. The study's results contribute to the growing body of knowledge in this field and have practical implications for improving patient care and outcomes in the transplantation setting.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Grant support:

This study was partially funded by the American Heart Association (AHA; grant number 19GPSGC35180031) and by the National Institute on Aging (NIH/NIA; grant number R01AG084236).

Abbreviations and Acronyms:

AUC-PR	area under the precision-recall curve
AUC-ROC	area under the receiver operating characteristic curve
BiGRU	bidirectional gated recurrent units
CAD	coronary artery disease
LT	liver transplantation
MACE	major adverse cardiovascular events
MASH	metabolic dysfunction-associated steatohepatitis

REFERENCES

1. VanWagner LB, Serper M, Kang R, et al. Factors associated with major adverse cardiovascular events after liver transplantation among a national sample. Am J Transplant. 2016;16(9):2684–2694. 10.1111/ajt.13779. [PubMed: 26946333]
2. Hughes DL, Rice JD, Burton JR, et al. Presence of any degree of coronary artery disease among liver transplant candidates is associated with increased rate of post-transplant major adverse cardiac events. Clin Transplant. 2020;34(11):e14077. 10.1111/ctr.14077. [PubMed: 32939833]
3. Cheng XS, VanWagner LB, Costa SP, et al. Emerging evidence on coronary heart disease screening in kidney and liver transplantation candidates: a scientific statement from the American Heart Association. Circulation. 2022;146(21):e299–e324. 10.1161/cir.0000000000001104. [PubMed: 36252095]
4. VanWagner LB, Harinstein ME, Runo JR, et al. Multidisciplinary approach to cardiac and pulmonary vascular disease risk assessment in liver transplantation: an evaluation of the evidence and consensus recommendations. Am J Transplant. 2018;18(1):30–42. 10.1111/ajt.14531. [PubMed: 28985025]
5. Barman PM, VanWagner LB. Cardiac risk assessment in liver transplant candidates: current controversies and future directions. Hepatology. 2021;73(6):2564–2576. 10.1002/hep.31647. [PubMed: 33219576]
6. VanWagner LB, Ning H, Whitsett M, et al. A point-based prediction model for cardiovascular risk in orthotopic liver transplantation: the CAR-OLT score. Hepatology. 2017;66(6):1968–1979. 10.1002/hep.29329. [PubMed: 28703300]
7. Koshy AN, Gow PJ, Han H-C, et al. Sudden cardiac death following liver transplantation: incidence, trends and risk predictors. Int J Cardiol. 2021;327:171–174. 10.1016/j.ijcard.2020.11.038. [PubMed: 33253736]
8. Levy PE, Khan SS, VanWagner LB. Cardiac evaluation of the kidney or liver transplant candidate. Curr Opin Organ Transplant. 2020;26(1):77–84. 10.1097/mot.0000000000000838.
9. Shroff H, VanWagner LB. Cardiovascular disease in nonalcoholic steatohepatitis: screening and management. Curr Hepatol Rep. 2020;19(3):315–326. 10.1007/s11901-020-00530-0. [PubMed: 33585157]
10. Koshy AN, Farouque O, Cailes B, et al. Prediction of perioperative cardiovascular events in liver transplantation. Transplantation. 2020;105(3):593–601. 10.1097/tp.0000000000003306.

11. Watt KDS, 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(6):1420–1427. 10.1111/j.1600-6143.2010.03126.x. [PubMed: 20486907]
12. Kutkut I, Rachwan RJ, Timsina LR, et al. Pre–liver transplant cardiac catheterization is associated with low rate of myocardial infarction and cardiac mortality. Hepatology. 2020;72(1):240–256. 10.1002/hep.31023. [PubMed: 31696952]
13. Optum Claims Data. Optum. https://www.optum.com/business/life-sciences/real-world-data/claims-data.html. Accessed February 10, 2024.
14. Bahdanau D, Cho K, Bengio Y. Neural Machine Translation by Jointly Learning to Align and Translate. Cornell University; 2015. https://arxiv.org/abs/1409.0473.
15. Sundararajan M, Taly A, Yan Q. Axiomatic attribution for deep networks. https://arxiv.org/abs/1703.01365. Accessed May 31, 2021.
16. Tran J, Sharma D, Gotlieb N, Xu W, Bhat M. Application of machine learning in Liver transplantation: a review. Hepatol Int. 2022;16(3):495–508. 10.1007/s12072-021-10291-7. [PubMed: 35020154]
17. Spann A, Yasodhara A, Kang J, et al. Applying machine learning in liver disease and transplantation: a comprehensive review. Hepatology. 2020;71(3):1093–1105. 10.1002/hep.31103. [PubMed: 31907954]
18. 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(1):107–114. 10.1097/01.tp.0000442781.31885.a2. [PubMed: 24621533]
19. Umphrey LG, Hurst RT, Eleid MF, et al. Preoperative dobutamine stress echocardiographic findings and subsequent short-term adverse cardiac events after orthotopic liver transplantation. Liver Transpl. 2008;14(6):886–892. 10.1002/lt.21495. [PubMed: 18508373]
20. VanWagner LB, Lapin B, Levitsky J, et al. High early cardiovascular mortality after liver transplantation. Liver Transpl. 2014;20(11):1306–1316. 10.1002/lt.23950. [PubMed: 25044256]
21. Nitski O, Azhie A, Qazi-Arisar FA, et al. Long-term mortality risk stratification of liver transplant recipients: real-time application of deep learning algorithms on longitudinal data. Lancet Digit Health. 2021;3(5):e295–e305. 10.1016/s2589-7500(21)00040-6. [PubMed: 33858815]
22. Delen D, Oztekin A, Kong ZJ. A machine learning-based approach to prognostic analysis of thoracic transplantations. Artif Intell Med. 2010;49(1):33–42. 10.1016/j.artmed.2010.01.002. [PubMed: 20153956]
23. Jain V, Bansal A, Radakovich N, et al. Machine learning models to predict major adverse cardiovascular events after orthotopic liver transplantation: a cohort study. J Cardiothorac Vasc Anesth. 2021;35(7):2063–2069. 10.1053/j.jvca.2021.02.006. [PubMed: 33750661]

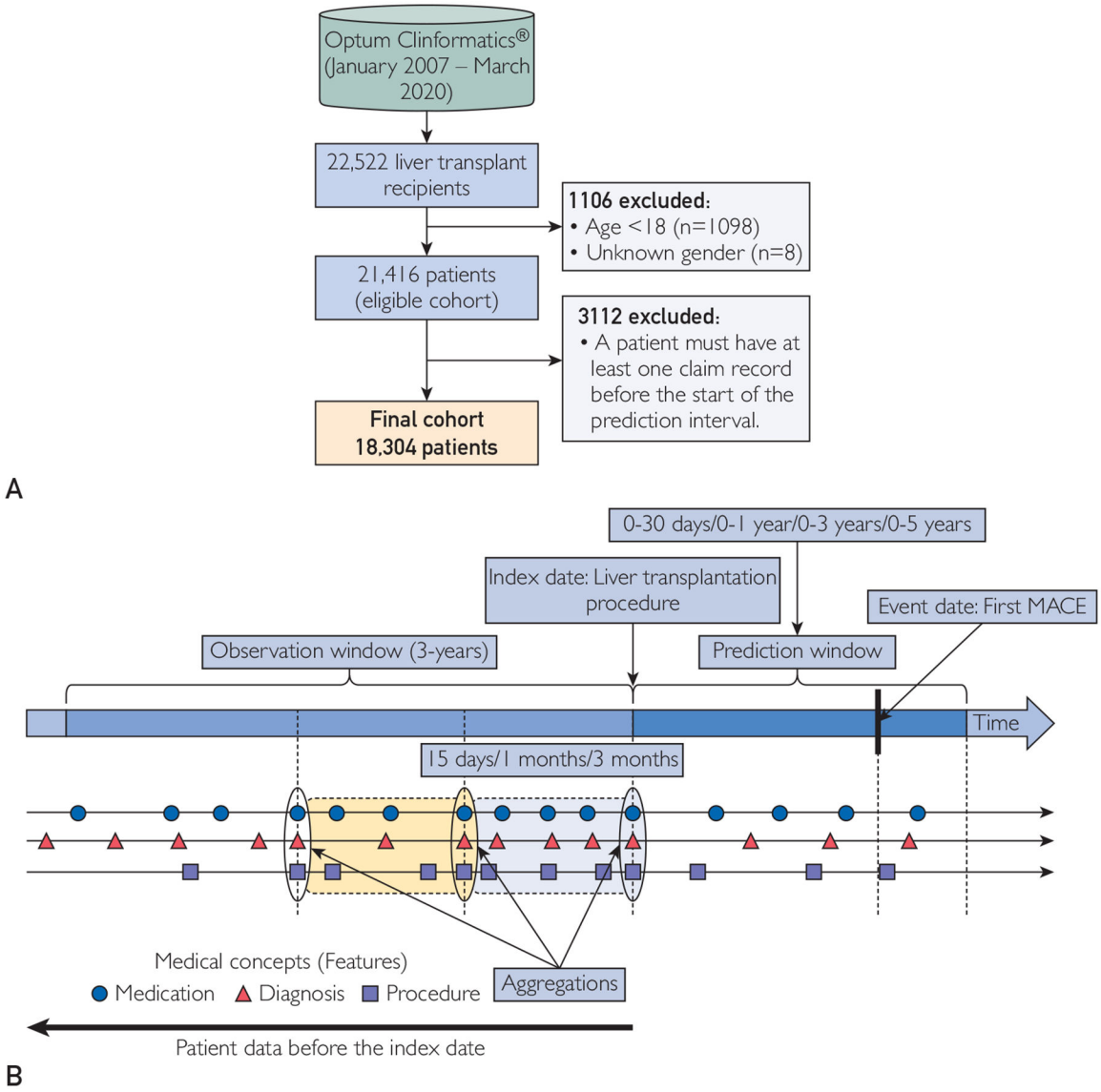


FIGURE 1.

Study design: (A) Cohort construction. (B) The study design diagram showing the aggregation of medical concepts before the index date. The study's objective is shown as well: to predict the first MACE within a defined prediction window after liver transplantation. MACE, major adverse cardiovascular event.

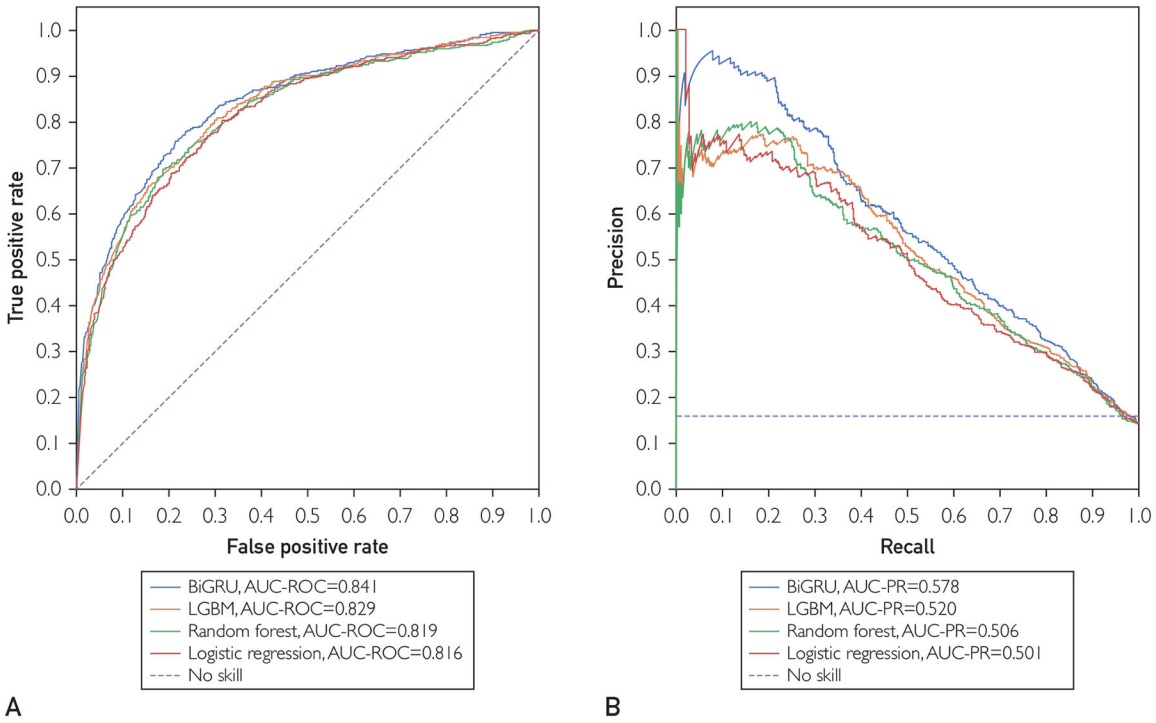


FIGURE 2.

Performance comparisons of the deep learning model and other machine learning models for major adverse cardiovascular event prediction in the 0- to 30-day interval. The blue line is used to show the performance of the bidirectional gated recurrent units (BiGRU) model; the orange line for the LGBM model; the green line for the random forest; and the red line for the logistic regression. (A) Comparison of the models' areas under the receiver operating characteristic curves (AUC-ROCs). (B) Comparison of the models' areas under the precision-recall curves (AUC-PR).

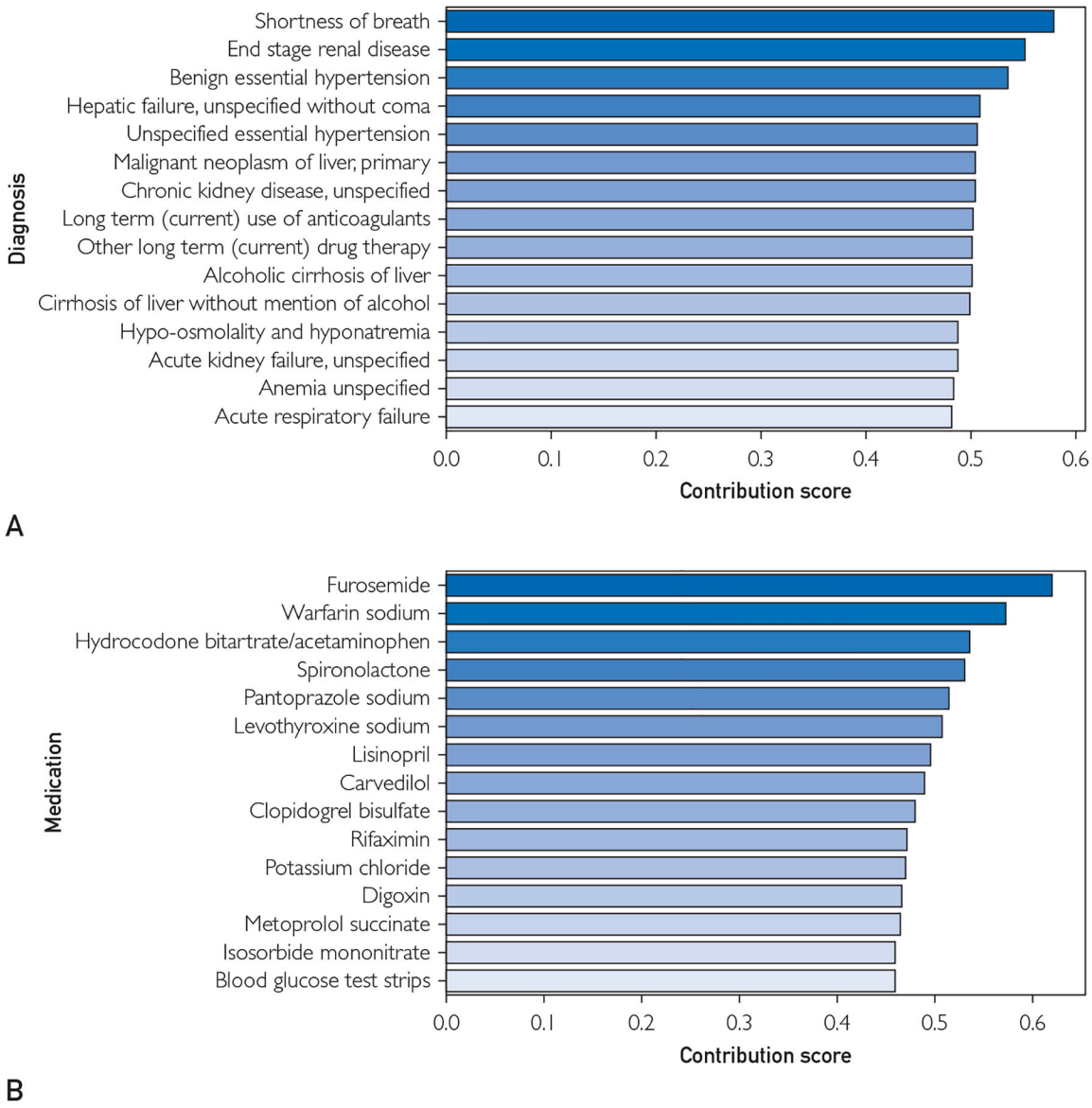


FIGURE 3.

Ranking by importance of the top 15 predictors (diagnoses and medications) for major adverse cardiovascular event prediction using the bidirectional gated recurrent units bidirectional gated recurrent unit deep learning model in the 0- to 30-day interval. (A) Top-ranked diagnoses. (B) Top-ranked medications.

TABLE 1.

Cohort demographic and clinical characteristics of patients before the LT index date

Characteristic	NO_MACE n=11,843 (64.7%)	MACE n=6461 (35.30%)	Overall n=18,304	*P*
Age (mean)	55 y	62 y	57 y	<.001[a]
Gender_F	4722 (39.9%)	2436 (37.7%)	7158 (39.1%)	.004
Gender_M	7121 (60.1%)	4025 (62.3%)	11,146 (60.9%)	.004
BMI	1011 (8.5%)	791 (12.2%)	1802 (9.8%)	<.001[a]
Smoking	2063 (17.4%)	1607 (24.9%)	3670 (20.1%)	<.001[a]
Diabetes	3294 (27.8%)	2881 (44.6%)	6175 (33.7%)	<.001[a]
History of CAD	1204 (10.2%)	1734 (26.8%)	2938 (16.1%)	<.001[a]
Hypertension	5096 (43.0%)	4049 (62.7%)	9145 (50.0%)	<.001[a]
Previous MI	808 (6.8%)	1213 (18.8%)	2021 (11.0%)	<.001[a]
PVD	444 (3.7%)	649 (10.0%)	1093 (6.0%)	<.001[a]
Dyslipidemia	3448 (29.1%)	2843 (44.0%)	6291 (34.4%)	<.001[a]
Previous stroke	435 (3.7%)	804 (12.4%)	1239 (6.8%)	<.001[a]
Pulmonary hypertension	511 (4.3%)	891 (13.8%)	1402 (7.7%)	<.001[a]
Alcohol abuse	1180 (10.0%)	759 (11.7%)	1939 (10.6%)	<.001[a]
MASH	1590 (13.4%)	1154 (17.9%)	2744 (15.0%)	<.001[a]

[a]*P* value: the χ^2 test was used for all fields except age, which was calculated by the Kruskal–Wallis test.

Percentages are based on the total number of patients in each group (MACE/NO_MACE).

BMI, body mass index; CAD, coronary artery disease; F, female; M, male; MI, myocardial infarction; PVD, peripheral vascular disease.

TABLE 2.

Comparison of model performance between the BiGRU deep learning model and 3 machine learning models using the AUC-ROC and AUC-PR

Model	Performance measure	Prediction scenarios			
		0–30 d (95% CI)	0–1 y (95% CI)	0–3 y (95% CI)	0–5 y (95% CI)
BiGRU	AUC-ROC	0.841 (0.822–0.862)	0.789 (0.772–0.806)	0.752 (0.737–0.767)	0.763 (0.748–0.777)
	AUC-PR	0.578 (0.536–0.621)	0.655 (0.623–0.683)	0.639 (0.613–0.661)	0.682 (0.661–0.711)
LGBM	AUC-ROC	0.829 (0.81–0.85)	0.784 (0.771–0.802)	0.750 (0.734–0.767)	0.762 (0.744–0.778)
	AUC-PR	0.520 (0.478–0.57)	0.638 (0.612–0.670)	0.646 (0.62–0.673)	0.601 (0.659–0.713)
Random forest	AUC-ROC	0.819 (0.793–0.837)	0.777 (0.759–0.795)	0.740 (0.724–0.757)	0.757 (0.740–0.772)
	AUC-PR	0.506 (0.455–0.555)	0.611 (0.58–0.639)	0.618 (0.593–0.645)	0.665 (0.638–0.689)
Logistic regression	AUC-ROC	0.816 (0.794–0.832)	0.767 (0.748–0.785)	0.742 (0.726–0.758)	0.756 (0.742–0.771)
	AUC-PR	0.500 (0.461–0.548)	0.598 (0.567–0.627)	0.628 (0.605–0.655)	0.666 (0.64–0.695)

AUC-PR, area under the precision-recall curve; AUC-ROC, area under the receiver operating characteristic curve; BiGRU, bidirectional gated recurrent units; LGBM, light gradient-boosting machine.