Received: 1 November 2021 | Revised: 3 March 2022 | Accepted: 8 March 2022

DOI: 10.1002/hep.32457

ORIGINAL ARTICLE

Accurate long-term prediction of death for patients with cirrhosis

David Goldberg[1,2] | **Alejandro Mantero**[2] | **David Kaplan**[3,4] | **Cindy Delgado**[1] | **Binu John**[1,5] | **Nadine Nuchovich**[1] | **Ezekiel Emanuel**[6] | **Peter P. Reese**[7]

[1]Division of Digestive Health and Liver Diseases, Department of Medicine, University of Miami Miller School of Medicine, Miami, Florida, USA

[2]Department of Public Health Sciences, University of Miami Miller School of Medicine, Miami, Florida, USA

[3]Division of Gastroenterology, Department of Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA

[4]Corporal Michael J. Crescenz VA Medical Center, Philadelphia, Pennsylvania, USA

[5]Bruce Carter VA Medical Center, Miami, Florida, USA

[6]Department of Medical Ethics and Health Policy, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA

[7]Renal-Electrolye and Hypertension Division, Department of Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania, USA

Correspondence
David Goldberg, Don Soffer Clinical Research Building, 1120 NW 14th Street, Room 807, Miami, FL 33136, USA.
Email: dsgoldberg@miami.edu

Funding information
Supported by the National Institutes of Health (R01 DK120561, to D. G.)

Abstract

Background and Aims: Cirrhosis is a major cause of death and is associated with extensive health care use. Patients with cirrhosis have complex treatment choices due to risks of morbidity and mortality. To optimally counsel and treat patients with cirrhosis requires tools to predict their longer-term liver-related survival. We sought to develop and validate a risk score to predict longer-term survival of patients with cirrhosis.

Approach and Results: We conducted a retrospective cohort study of adults with cirrhosis with no major life-limiting comorbidities. Adults with cirrhosis within the Veterans Health Administration were used for model training and internal validation, and external validation used the OneFlorida Clinical Research Consortium. We used four model-building approaches including variables predictive of cirrhosis-related mortality, focused on discrimination at key time points (1, 3, 5, and 10 years). Among 30,263 patients with cirrhosis ≤75 years old without major life-limiting comorbidities and complete laboratory data during the baseline period, the boosted survival tree models had the highest discrimination, with 1-year, 3-year, 5-year, and 10-year survival rates of 0.77, 0.81, 0.84, and 0.88, respectively. The 1-year, 3-year, and 5-year discrimination was nearly identical in external validation. Secondary analyses with imputation of missing data and subgroups by etiology of liver disease had similar results to the primary model.

Conclusions: We developed and validated (internally and externally) a risk score to predict longer-term survival of patients with cirrhosis. This score would transform management of patients with cirrhosis in terms of referral to specialty care and treatment decision-making for non-liver-related care.

Abbreviations: eGFR, estimated glomerular filtration rate; HFrEF, heart failure with a reduced ejection fraction; ICD-9, *International Classification of Diseases*, Ninth Revision; INR, international normalized ratio; IQR, interquartile range; LT, liver transplantation; LTCS, long-term cirrhosis survival; MELD, Model for End-Stage Liver Disease; PH, proportional hazards; VA, Veterans Administration; VHA, Veterans Health Administration; VIMP, variable importance.

INTRODUCTION

In 2017, there were 1.32 million worldwide deaths from cirrhosis, representing 2.4% of all deaths.[1] Cirrhosis-related mortality is increasing worldwide (44% increase in the last 20 years) and in the United States (65% increase from 1999 to 2016), most notably in older patients with NASH.[1–4] For clinicians, the ability to predict longer-term survival of patients with cirrhosis is important to counsel them, especially in the face of complex treatment choices that may be related to their liver disease (e.g., transplantation) or comorbidity management where liver-related mortality may compromise outcomes and/or obviate the benefits of treatment. However, estimating longer-term liver-related survival for patients with cirrhosis with a high degree of accuracy remains a challenge because the focus of cirrhosis risk prediction has been short-term survival (i.e., 90 days), the basis for waitlist priority in the United States and many other countries.

The Model for End-Stage Liver Disease (MELD) score was developed over two decades ago to predict short-term survival among patients waitlisted for transplant.[5–7] Despite refinements, it has poor discrimination among patients with low scores or as a tool to predict long-term survival.[5–9] Although models have been developed to predict longer-term survival of patients with cirrhosis, each has important limitations. For example, a model developed using machine learning in the Veterans Health Administration (VHA) (1) had a decrement in performance beyond 1 year and (2) included patients with life-limiting conditions (e.g., metastatic cancer, heart failure) that may have been just as likely to cause death compared to cirrhosis, thereby preventing the score from predicting liver-specific mortality.[10,11] Another model used the MELD score and two-dimensional shear wave elastography but was limited by not excluding patients with major life-limiting medical comorbidities; performed best categorizing patients as having “good,” “intermediate,” or “poor” prognosis; and focused only over a 2-year horizon.[12]

The ability to accurately predict cirrhosis-related mortality could provide tremendous benefits to patients and their surrogates. First, accurate survival prediction could improve timing of referral for advanced cirrhosis care by identifying patients who would benefit from closer management by hepatologists and a multidisciplinary care team. Second, it could help patients and caregivers make health care decisions armed with accurate data about expected outcomes, especially with advanced care planning where limitations in survival prediction beyond the short term are a barrier.[13] Third, it could help guide clinical management where liver-related mortality may compromise outcomes and/or obviate benefits of treatment (e.g., cancer therapy, surgical vs. medical management of cardiovascular disease). As the population with cirrhosis ages, these complex scenarios will become more common, especially because transplant is not an option for most patients.[14,15] We therefore sought to develop and validate (internally and externally) long-term risk scores among patients with cirrhosis without other life-limiting medical comorbidities (long-term cirrhosis survival [LTCS]).

PATIENTS AND METHODS

Data source for model development and internal validation

We conducted a retrospective cohort study among patients with cirrhosis in the VHA using an updated cohort (i.e., eligibility period and follow-up) of patients in the Veterans Outcomes and Costs Associated with Liver Disease study group.[16–21] The VHA was selected because it (1) is the largest provider of liver care in the United States; (2) has a comprehensive electronic medical record that includes inpatient and outpatient clinical, laboratory, imaging, and prescription data tracked across all VHA facilities; (3) has complete death data ascertainment; and (4) has a diverse population that reflects the racial/ethnic distribution of the United States. Because only a fraction of patients with cirrhosis in the VHA are waitlisted for liver transplantation (LT),[22–24] in contrast to transplant registry data where >50% of patients are transplanted within 1 year of waitlisting,[25,26] the VHA provides a unique opportunity to model the natural history of cirrhosis among a diverse cohort.

Data collection

Demographic, clinical, laboratory, imaging, and administrative coding data from all VHA sites were obtained from the VHA relational database, cancer data from the VHA Tumor Registry, and mortality data from the Vital Status File after approval from the institutional review boards at the University of Miami, the Miami Veterans Administration (VA) Health System, and the Corporal Michael J. Crescenz VA Medical Center.

Cohort selection

We included adults aged ≥18 years and ≤75 years at cirrhosis diagnosis between January 1, 2008, and December 28, 2018. Time zero (start of follow-up) was at cirrhosis diagnosis. The upper age limit was chosen to generalize to the LT waitlist population because <0.3% of waitlist additions are >75 years old.[27] Cirrhosis was defined using validated methods of one inpatient and/or two outpatient *International Classification of Diseases*, Ninth Revision (ICD-9; 571.2 or 571.5) or ICD-10 (K74.60, K74. 69, K70.30, K70.31) codes.[16–21]

We restricted analyses to patients actively engaged in care within the VHA (i.e., ≥2 outpatient visits) from January 1, 2008, to December 28, 2018.[16–21] We excluded patients with HCC (based on diagnosis codes and VHA tumor registry) prior to their cirrhosis diagnosis, and censored at HCC diagnosis after the index date because predictors of survival in these patients include tumor-specific variables that do not apply to the general cirrhosis population and are beyond the scope of this work.

We excluded patients with non-liver life-limiting comorbidities that could be reliably ascertained using electronic medical record data. First, this allowed us to better focus on prediction of liver-related mortality by excluding conditions that would be a competing risk to liver-related mortality. Second, we excluded patients with absolute contraindications to LT to ensure that the score could be used to predict cirrhosis-related mortality and long-term survival benefit of transplant for potential transplant candidates. The specific criteria are listed in Table S1, but the exclusions were (1) heart failure with a reduced ejection fraction (HFrEF) by echocardiogram, (2) severe obstructive/restrictive lung disease, (3) nonlocalized cancer (i.e., regional, lymph node, metastatic) or localized cancer (excluding nonmelanoma skin cancer and prostate cancer) diagnosed within 5 years of cirrhosis index date, and/or (4) uncontrolled HIV/AIDS.[28] The HFrEF exclusion required an abnormal echocardiogram, but patients with no echocardiogram were included because >80% of patients with HFrEF are symptomatic and would be expected to have an echocardiogram.[29,30] We applied broad criteria for cancer exclusions to exclude (1) nonliver cancer as a cause of mortality (i.e., competing risk to liver disease–related mortality) and (2) any cancer diagnosis that could contraindicate a transplant to generalize to the waitlist population.

Covariates

Models included baseline demographic (e.g., age), clinical (e.g., diabetes, complications of cirrhosis rather than decompensation stages, diabetes), and laboratory (e.g., international normalized ratio [INR], renal function) variables associated with short-term and/or long-term mortality in patients with cirrhosis. The time window to collect baseline data was −180 days to +90 days from the cirrhosis index date, with the closest date selected when multiple laboratory values were available. This time window reflects the standard approach in cohort studies using administrative data and has face validity that baseline data reflected the patient's clinical status at cirrhosis diagnosis (i.e., rather than 1 year prediagnosis). Renal function was modeled as (1) continuous estimated glomerular filtration rate (eGFR), using the Modification of Diet in Renal Disease-4 equation[31–33]; (2) acute kidney injury (creatinine increase of ≥0.3 mg/dl in ≤48 h and/or increase of >50% in 7 days)[34–36]; and (3) chronic kidney disease (Kidney Disease: Improving Global Outcomes classification of eGFR < 60 ml/min/1.73 m^2 at every time point over a 90-day period).[31–33] These methods allowed us to quantify the degree of renal impairment using a continuous value (eGFR) and, when abnormal, whether it was acute, chronic, or acute on chronic kidney injury. Complications of cirrhosis at diagnosis were ascertained using validated methods.[16–21]

Statistical analysis

The primary outcome was overall survival, with models assessing discrimination at key time points, given the inability to specifically define liver-related versus liver-unrelated death.[2,37] Time-dependent receiver operating characteristic curves were constructed, and AUCs were compared at the 1-year, 3-year, 5-year, and 10-year time horizons to determine comparative model performance. Specifically, we used the method of cumulative sensitivity/dynamic specificity and truncated follow-up at 10 years (i.e., censored outcomes after 10 years) due to the clinical focus of the model and the limited follow-up beyond this time point. We tested four model-building methodologies with the same possible covariates: (1) standard regression (Cox proportional hazards [PH]), (2) stepwise Cox PH, (3) machine-learning boosted survival trees, and (4) efficient neural networks.[11,38,39] These methods were used because many patients had only one respective lab value during the baseline. For model development, we used an 80/20 training/validation split, the standard method of model derivation/validation to maximize observations for training while providing sufficient sample size to provide narrow CIs in the validation.[40–43] Model calibration was performed in the VA training set.

For standard Cox PH models, we selected variables until the model's AUC increased by <0.001[44] For stepwise Cox PH, we used the stepAIC function from the MASS package where variables are selected and deselected to minimize the Akaike information criterion. For boosted survival trees, we used a standard approach whereby variables were trained on the outcome using a tree-based algorithm to create a set of shallow trees with the glmnet package that could not be more than a few levels deep, were iteratively selected and included by "slow learning," and repeated for 1000 iterations with 10-fold cross-validation for calibrating the optimal iteration within the training set; additionally, the out of bag estimates were used in this calibration in order to best characterize test set optimization.[45–47] These boosted survival trees are a type of ensemble method using more than one decision tree in the data mining process (i.e., in contrast to classification and

regression tree analysis that uses one decision tree). For efficient neural networks, variables were allowed to be partially in and out of the model through "slow learning" using the gbm package, with iterative model training to prevent overfitting of the data.[48,49] Elastic net Cox PH is a form of constrained optimization based on Cox regression where the idea is to "loosen" the fit in order to reduce overfitting on the training data with the aim of increasing predictive performance.

For the final model using boosted survival trees (results), there were no coefficients for individual variables (compared to beta coefficients from a logistic regression model) by virtue of it being a tree-based model.[45–47] Therefore, we plotted relative variable importances (VIMPs) for the model's predictors, obtained by scaling the VIMP values from 0 to 1 by dividing by the maximum importance.[45,50] VIMP is obtained by randomly permuting each predictor and calculating the increase in error produced.[50] The higher the increase, the bigger a role that variable played in correctly predicting cases (converse is true). This metric is usually used when we have "black box" models where the role of predictors cannot be easily summarized directly from the trained model.[45,50] All models censored (competing risk models not possible using machine learning methods) patients at LT given that this interrupted the "natural history" of their cirrhosis (only 2.5% of the final cohort received a transplant during follow-up).

We selected the final model based on the highest discrimination (i.e., highest AUC truncating survival at 10 years using the method described by Chambless and Diao[51]) and converted the output into a clinically interpretable risk score by (1) multiplying XBETA by −1 so that higher scores were "better" (longer survival) and (2) multiplying the score by 10 to generate a scale comparable to the MELD score scale familiar to clinicians. Although the C-statistics reflect the score's discrimination as a continuous variable, we categorized patients into quartiles of LTCS score to demonstrate the score's ability to stratify patients by actual survival. To graphically present data, we categorized the testing cohort into four groups based on quartiles of their LTCS score to graph survival curves and to calculate the mean survival within each quartile. AUC values are accompanied by 95% CIs that were empirically derived through 20 replicates because (1) computation time for a cohort this size meant about 5 min per iteration and (2) at 20 replicates the CI was already very narrow due to the large sample size.

In primary analyses we performed complete-case analysis because the (1) data were not missing completely at random; (2) approach is consistent with methods used to develop other organ allocation scores (e.g., MELD[5-7], Lung Allocation score); and (3) number of cases missing ≥1 data point was large (>50%), which may lead to unstable imputation with suboptimal performance.[52–54] However, to address potential biases, we performed a sensitivity analysis using imputation through the missForest package among patients missing one or two laboratory values to maximize the imputation's performance while not imputing too many variables.[52–55] After selecting the final model, we estimated survival time for a given LTCS score by multiplying the estimated baseline survival function for the entire population by the specific conditional estimated risk based on the boosting algorithm. This allowed us to calculate the overall estimated survival function for an individual, for which we could then determine the time at which we expected median survival and, therefore, reported predicted survival time.[51] Lastly, we calculated the Brier score of the final model to evaluate the accuracy of outcome predictions from our model[56] and fit calibration plots using a gradient-boosted, machine-based function to assess the agreement between the predicted and observed survival.[57]

External validation

We performed external validation using the OneFlorida Clinical Research Consortium, a statewide database of clinical and laboratory data from 11 health care systems and affiliated practices that provide health care to >40% of Floridians, reaching 15 million patients since its inception in 2013.[58–60] The OneFlorida Consortium includes demographic, clinical (ICD-9/10 codes), medication, and laboratory data.[58–60] Therefore, it was a resource to externally validate the LTCS score in a diverse cohort of patients with available laboratory data. We identified patients with cirrhosis based on ICD-9 codes diagnosed on or after October 1, 2015 (introduction of ICD-10 codes), and applied similar inclusions/exclusions as in our main analysis. However, because OneFlorida does not have the same data as the VHA (e.g., echocardiograms), we had to rely only on ICD-10 codes for exclusionary diagnoses (e.g., HFrEF).

Secondary analyses

Although the MELD and MELD-Na scores were developed to predict 90-day mortality, we compared the performance of our four models to the MELD and MELD-Na scores over longer-term intervals and the final model to MELD and MELD-Na over a 90-day time horizon. We then fit models using the four variables in the MELD-Na score (sodium, INR, bilirubin, and creatinine) to our VHA cohort to compare the performance of a score using these variables (with their respective beta coefficients fit based on the cohort) to the LTCS score. Second, we stratified the performance of the final model based on etiology of cirrhosis. Third, we compared the performance of the LTCS score to the Fibrosis-4 (FIB-4) score, the aspartate

aminotransferase to platelet ratio index (APRI) score, and platelet count. Lastly, we assessed the final model's performance in patients with "less advanced" forms of cirrhosis, defined three ways: (1) calculated MELD score < 10, (2) Child-Turcotte-Pugh class A, and (3) compensated cirrhosis.

RESULTS

From January 1, 2008, to December 28, 2018, there were 163,008 veterans with ≥2 outpatient visits diagnosed with cirrhosis, of whom 74,997 (46.0%) were excluded for age >75 years or a life-limiting medical comorbidity (heart failure exclusions, 27,040; AIDS/uncontrolled HIV, 1394; severe lung disease, 23,929; cancer, 8200; Figure S1). Of the remaining 88,011 patients, 30,263 (34.3%) had complete laboratory data during the baseline; of 56,065 patients with incomplete laboratory data, 20,765 were missing one value, while 35,300 were missing ≥2 laboratory values. Among these 88,011 patients, the median age at cirrhosis diagnosis was 60.1 years (interquartile range [IQR], 55.4–64.3), 81,845 (96.9%) were male, 54,492 (64.5%) were White non-Hispanic, 14,418 (17.1%) were Black non-Hispanic, 7,370 (8.7%) were White Hispanic, >75% had alcohol-related and/or HCV-related liver disease, and 23,171 (27.4%) and 15,760 (18.7%) had ascites and HE at cirrhosis diagnosis, respectively, similar to other published cohorts.[4] Those with complete laboratory data were objectively sicker with higher mortality/lower survival (Table 1). The median follow-up was 1586 days (IQR, 864.25–2641 days), and the median time to event was 653 days (IQR, 211–1399.75 days).

Model output for risk scores

The boosted survival tree had the highest discrimination, and the model using missForest imputation had similar, but slightly lower, discrimination as the complete-case analysis (Table 2). The model had excellent prediction with a Brier score of 0.06, with calibration plots showing excellent calibration (Figure S2A–D). Of the variables in the boosted survival tree model, albumin was the strongest predictor of survival, followed by bilirubin, INR, and ascites, while the interactions of eGFR × acute kidney injury and chronic kidney disease were the weakest (Figure 1).

There was a separation in survival within the first year after cirrhosis diagnosis, with a nearly 40% difference in the mortality rate between patients in the top and bottom quartiles at 1 year, which increased to nearly 60% at 3 years (Figure 2A; Table S2). The median survival time ranged from 1.95 years (lowest quartile) to 12.0 years (highest quartile), with some overlap in the IQRs (Figure 2B).

External validation

The OneFlorida Clinical Research Consortium included 31,720 patients with cirrhosis, of whom 17,127 met inclusion/exclusion criteria and 7147 (41.7%) had full laboratory data. Their clinical and laboratory data were similar to those of the VHA cohort (Table S3), with the exception of nearly 40% of patients being female and higher 1-year, 3-year, and 5-year survival rates. The AUC of the model at the 1-year, 3-year, and 5-year time points was nearly identical to that of the VHA cohort (Table 2).

Secondary analyses

At the 1-year time horizon, the MELD-Na score had similar discrimination to the LTCS score, but while the MELD-Na discrimination decreased over time, the boosted survival tree LTCS model improved (Table 3; performance using the MELD-Na components with recalibrated beta coefficients was similar to that of the original MELD-Na). Furthermore, the LTCS score had superior discrimination at 90 days (Table 3). The final LTCS model had excellent discrimination across all disease entities (Table 4). There was a slight decrement in performance in patients with a baseline MELD score < 10 or compensated cirrhosis, although the performance at 10 years was excellent, with similar performance in patients with Child-Turcotte-Pugh class A cirrhosis (Table 5). The LTCS model performed substantially better than the APRI and FIB-4 scores, as well as a model only using baseline platelet count (Table S4). An interactive version of the LTCS calculator is available online at: https://amantero.shinyapps.io/plites/.

DISCUSSION

Using machine learning methods and >10 years of data, we developed a model with excellent ability to predict 1-year, 3-year, 5-year, and 10-year survival of patients with cirrhosis. Not only did we develop and internally validate the risk score but we externally validated it in a distinct cohort that generalizes to the broader population of patients with cirrhosis. This risk score can improve clinical care by better prognosticating patient survival, counseling patients and their close associates on treatment options and advanced care planning with better information, helping to select patients to refer for advanced cirrhosis care, and potentially in the future, allocating donor livers in a manner that optimizes the benefit of a scarce resource by considering longer-term survival benefit.

The LTCS score advances the field by predicting long-term outcomes among patients with cirrhosis with high discrimination using an externally validated

TABLE 1 Clinical and laboratory data of patients with cirrhosis ≤75 years of age without contraindications to LT[a]

Variable at time of diagnosis of cirrhosis	Analytic cohort (*n* = 30,263)	Excluded cohort (*n* = 57,748)[b]
Male gender, *n* (%)	29,354 (97.0)	55,976 (96.9)
Age, median (IQR), years	60.1 (55.4–64.2)	60.2 (20.8–75.0)
Race/ethnicity, *n* (%)		
White non-Hispanic	19,509 (64.5)	37,254 (64.5)
Black non-Hispanic	5217 (17.2)	9854 (17.1)
White Hispanic	2777 (9.2)	4921 (8.5)
Pacific Islander	229 (0.8)	463 (0.8)
Asian	120 (0.4)	189 (0.3)
Black Hispanic	61 (0.2)	180 (0.3)
Other	1138 (3.8)	2777 (4.8)
Unknown	1212 (4.0)	2110 (3.7)
Etiology of liver disease, *n* (%)		
EtOH-related liver disease only	9971 (32.9)	19,982 (34.6)
HCV + EtOH-related liver disease	9101 (30.1)	13,482 (23.3)
HCV only	5439 (18.0)	8746 (15.1)
NASH	2635 (8.7)	6113 (10.6)
Cryptogenic	1883 (6.2)	6972 (12.1)
HBV	651 (2.2)	1086 (1.9)
Hemochromatosis	271 (0.9)	557 (1.0)
PBC	137 (0.5)	357 (0.6)
AIH	93 (0.3)	229 (0.4)
PSC	82 (0.3)	224 (0.4)
Ascites, *n* (%)	10,383 (34.3)	13,748 (23.8)
Encephalopathy, *n* (%)	6716 (22.2)	9683 (16.8)
SBP, *n* (%)	1916 (6.3)	2697 (4.7)
Diabetes, *n* (%)	3309 (10.9)	2005 (3.9)
CKD, *n* (%)[c]	4705 (15.5)	5178 (9.0)
eGFR, median (IQR), mL/min/1.73 m^2 [c]	90.8 (71.0–110.0)	88.4 (71.4–107)
Sodium, median (IQR), mEq/L[c]	138 (135–140)	138 (136–140)
Albumin, median (IQR), g/L[c]	3.4 (2.7–3.9)	3.6 (3.0–4.0)
INR, median (IQR)[c]	1.2 (1.1–1.4)	1.17 (1.0–1.4)
Total bilirubin, median (IQR), mg/dL[c]	1.1 (0.7–2.1)	1.0 (1.1–1.4)
AST, median (IQR), units/L[c]	60 (36–100)	52 (32.0–89.0)
ALT, median (IQR), units/L[c]	43 (27–73)	41 (26–70)
AST/ALT ratio, median (IQR)[c]	1.3 (1.0–1.9)	1.3 (0.9–1.7)
Alkaline phosphatase, median (IQR), IU/L[c]	106 (78–150)	100 (75–140)
Hemoglobin, median (IQR), g/dl[c]	13.1 (11.2–14.6)	13.5 (11.8–14.9)
Platelet count, median (IQR), $10^3/\mu l$[c]	125 (85–181)	133 (90–191)
Unadjusted survival		
1-year	83.1% (82.6–83.5%)	87.4% (87.2–87.7%)
3-year	65.3% (64.7–65.8%)	71.0% (70.6–71.4%)
5-year	50.8% (50.2–51.4%)	58.2% (57.7–58.6%)
10-year	27.4% (26.6–28.1%)	36.8% (36.3–37.4%)

Abbreviations: AIH, autoimmune hepatitis; ALT, alanine aminotransferase; AST, aspartate aminotransferase; CKD, chronic kidney disease; EtOH, ethyl alcohol; PBC, primary biliary cholangitis; PSC, primary sclerosing cholangitis; SBP, spontaneous bacterial peritonitis.

[a]The analytic and excluded cohorts were statistically significantly different ($p < 0.05$) for all comparisons except for gender ($p = 0.61$) and chronic kidney disease ($p = 0.12$).

[b]Includes patients with insufficient lab data and/or death date concurrent with a diagnosis of cirrhosis.

[c]Lab data presented for patients in the excluded cohort, when available.

TABLE 2 Model discrimination for mortality prediction for patients with cirrhosis using four model-building approaches[a]

Model	1-year AUC	3-year AUC	5-year AUC	10-year AUC
Candidate risk models in internal validation cohort[b]				
Boosted survival trees	0.77 (0.77–0.77)	0.81 (0.81–0.82)	0.84 (0.83–0.84)	0.88 (0.87–0.88)
Efficient neural network	0.70 (0.69–0.70)	0.71 (0.71–0.72)	0.73 (0.73–0.73)	0.76 (0.76–0.77)
Cox PH	0.72 (0.72–0.73)	0.74 (0.74–0.74)	0.76 (0.75–0.76)	0.79 (0.79–0.79)
Stepwise Cox PH	0.72 (0.72–0.73)	0.74 (0.74–0.74)	0.76 (0.75–0.76)	0.79 (0.79–0.79)
MELD-Na (traditional model)	0.75 (0.72–0.78)	0.68 (0.68–0.69)	0.70 (0.69–0.70)	0.72 (0.72–0.73)
Performance of boosted survival tree model using imputation[c]	0.75 (0.75–0.76)	0.77 (0.77–0.77)	0.78 (0.78–0.78)	0.81 (0.81–0.81)
External validation of boosted survival tree model in OneFlorida[d]	0.77 (0.77–0.78)	0.80 (0.80–0.81)	0.83 (0.82–0.83)	N/A

Abbreviations: N/A, not available.

[a]The variables in the final model (boosted survival trees) with the highest predictive performance included age, baseline eGFR, chronic kidney disease (yes/no), sodium, diabetes mellitus (yes/no), HE, ascites, spontaneous bacterial peritonitis, albumin, INR, bilirubin, acute kidney injury, aspartate aminotransferase, alanine aminotransferase, aspartate aminotransferase/alanine aminotransferase ratio, alkaline phosphatase, hemoglobin, platelet count, interaction of eGFR × acute kidney injury. The 95% CIs were obtained by rerunning the analyses using 20 replicate bootstraps. Beta coefficients are not available for this type of model.

[b]Complete-case analysis included 30,263 patients.

[c]The multiple imputation included 63,487 patients.

[d]Based on external validation of 7147 patients with cirrhosis meeting inclusion/exclusion criteria with full laboratory data. Earliest index date of cirrhosis was October 1, 2015, so model performance capped at 5 years.

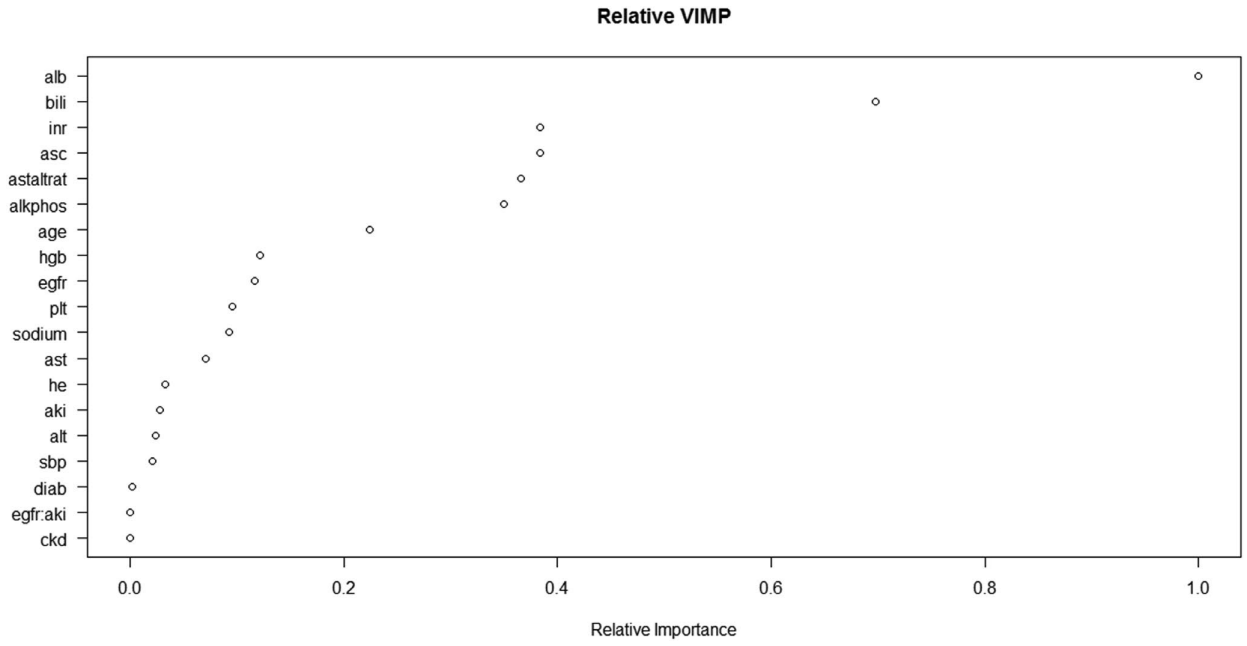

FIGURE 1 Relative VIMP plot for predictors included in the boosted survival model. Relative VIMPs obtained by scaling the VIMP values from 0 to 1 by dividing by the maximum importance. VIMP is obtained by randomly permuting each predictor and calculating the increase in error produced. The higher the increase, the bigger a role that variable played in correctly predicting cases.[45,50] Abbreviations: aki, acute kidney injury; alb, albumin; alkphos, alkaline phosphatase; alt, alanine aminotransferase; asc, ascites; ast, aspartate aminotransferase; astaltrat, aspartate aminotransferase to alanine aminotransferase ratio; bili, bilirubin; ckd, chronic kidney disease; diab, diabetes; egfr, EGF receptor; hgb, hemoglobin B; plt, platelets; sbp, spontaneous bacterial peritonitis

risk score. In contrast to studies using transplant registry data, we used VHA data to model the "natural history" of cirrhosis without being interrupted by LT.[27] LTCS has several features that make it a superior alternative to the score developed by the Baylor group using VHA data[11]: (1) higher discrimination (i.e., AUC) that improves over longer follow-up, 2) generalizability to a transplant/transplant-eligible population due to

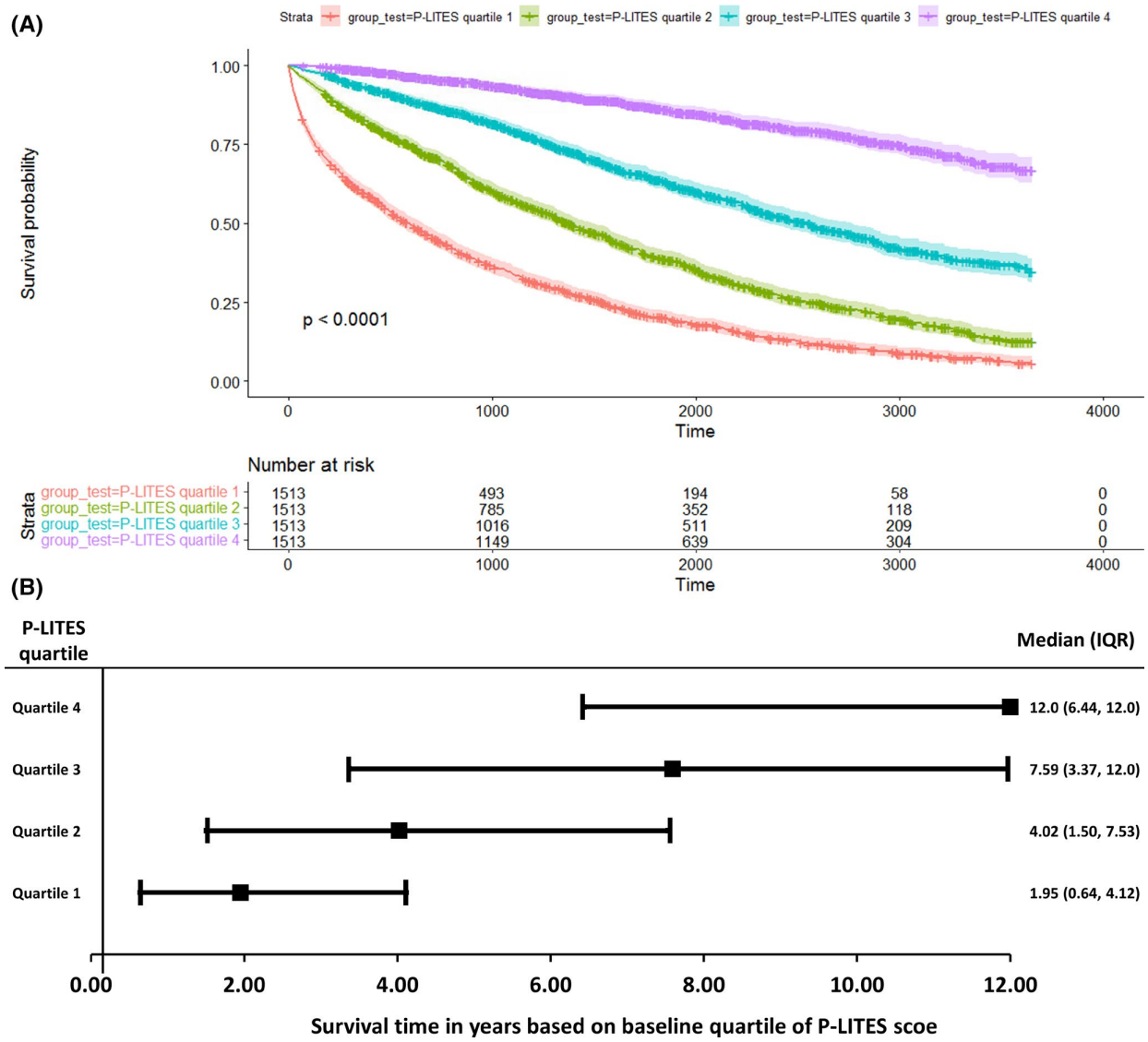

FIGURE 2 Survival of VHA validation cohort based on quartile of LTCS scores. (A) Survival probabilities of validation cohort based on quartile of LTCS score. (B) Absolute survival of validation cohort based on quartile of LTCS score. Quartiles of LTCS score are from the validation cohort of patients in the VHA

TABLE 3 Model discrimination for mortality prediction for patients with cirrhosis using boosted survival trees versus MELD-Na in the 20% VA testing cohort[a]

	Boosted survival trees AUC	MELD-Na AUC
3-month	0.74 (0.74–0.75)	0.67 (0.66–0.67)
1-year	0.77 (0.77–0.77)	0.75 (0.72–0.78)
3-year	0.81 (0.81–0.82)	0.68 (0.68–0.69)
5-year	0.84 (0.83–0.84)	0.70 (0.69–0.70)
10-year	0.88 (0.87–0.88)	0.72 (0.72–0.73)

[a]*n* = 6053 for the 20% sample.

age and comorbidity exclusions, (3) external validation in a population that was 40% female, and (4) greater focus on liver-related mortality by not including a co-morbidity index. The latter point is underscored by the fact that the covariate with the highest point estimate in the Baylor score was a cirrhosis-specific comorbidity score, demonstrating how its performance relies in part on predicting mortality from non-liver-related comorbidities. Additional strengths of our study are that we had a large sample size, assessed model discrimination imputation of missing data in addition to complete-case analysis, and performed external

TABLE 4 Model discrimination for mortality prediction for patients with cirrhosis based on etiology of liver disease using final boosted survival tree model in the 20% VA testing cohort[a]

Model	1-year AUC	3-year AUC	5-year AUC	10-year AUC
All patients	0.77 (0.77–0.77)	0.81 (0.81–0.82)	0.84 (0.83–0.84)	0.88 (0.87–0.88)
Alcohol-associated liver disease	0.77 (0.76–0.78)	0.82 (0.82–0.83)	0.86 (0.86–0.87)	0.91 (0.90–0.91)
Hepatitis C	0.76 (0.75–0.76)	0.78 (0.77–0.79)	0.80 (0.79–0.80)	0.83 (0.82–0.83)
NASH	0.76 (0.75–0.77)	0.79 (0.78–0.81)	0.82 (0.80–0.84)	0.87 (0.85–0.89)
Autoimmune hepatitis	0.74 (0.66–0.82)	0.78 (0.68–0.87)	0.80 (0.70–0.89)	0.82 (0.71–0.93)
Cholestatic liver disease	0.74 (0.65–0.83)	0.78 (0.68–0.88)	0.81 (0.73–0.90)	0.88 (0.80–0.96)

[a]Prediction accuracy based on the AUC for patients in the VHA test data set based on etiology of liver disease for the 20% testing sample of 6053 patients

TABLE 5 Secondary analyses assessing model discrimination for mortality prediction for specific subgroups of patients with cirrhosis[a]

Model	1-year AUC	3-year AUC	5-year AUC	10-year AUC
All patients	0.77 (0.77–0.77)	0.81 (0.81–0.82)	0.84 (0.83–0.84)	0.88 (0.87–0.88)
Baseline MELD score < 10	0.73 (0.73–0.74)	0.77 (0.76–0.77)	0.79 (0.78–0.80)	0.83 (0.84–0.84)
Child-Turcotte-Pugh class A	0.71 (0.70–0.72)	0.78 (0.77–0.79)	0.84 (0.83–0.85)	0.91 (0.91–0.93)
Compensated cirrhosis[b]	0.72 (0.71–0.72)	0.73 (0.73–0.74)	0.75 (0.74–0.75)	0.78 (0.78–0.79)

[a]Prediction accuracy based on the AUC for patients in the VHA test data set based on etiology of liver disease.
[b]Excluded patients with a history of a hepatic decompensation event (e.g., ascites) at baseline.

validation, thereby overcoming the most common methodological limitations of studies that use machine learning to develop risk prediction models.[61]

The ability to accurately predict cirrhosis-related mortality could provide tremendous benefits to patients and their surrogates. First, accurate survival prediction could improve timing of referral to a tertiary care center (e.g., transplant center) for advanced cirrhosis care.[14,15] Although multidisciplinary hepatology care is associated with lower mortality rates for patients with cirrhosis,[62] referral to a specialty/transplant center is often not considered outside of context transplantation.[14,15] Providers frequently use the MELD score as a trigger to refer a patient for advanced cirrhosis care, even though some patients with low MELD scores are at a substantial risk of death.[5–9] Use of the LTCS score could therefore improve the care of patients with cirrhosis and help to mitigate racial disparities in access to specialty cirrhosis care.[63] For example, a patient with a MELD score of 10 could have an LTCS score that confers a 3-year mortality of almost 50%. In the context of transplantation, such a patient could be advised to consider a living donor liver transplant given that it may be associated with superior survival over a longer-term time horizon (e.g., 3–5 years) and allow for transplantation before the patient develops complications (e.g., sarcopenia, frailty, chronic kidney disease) that may compromise long-term posttransplant outcomes.[31,32] Outside of transplantation, this patient could benefit from advanced hepatology care given an expected 5-year survival rate that is lower than that of regional colorectal cancer. And although organ allocation currently focuses on a "sickest first" approach based on short-term waitlist mortality, if in the future an integrated survival benefit–based approach that accounted for longer-term pretransplant and posttransplant survival probability were implemented, then the LTCS score could be integrated with the previously developed LITES scores (i.e., posttransplant survival).[31,32] Such an approach provides a superior population-based approach to allocation that ensures maximal life-years gained from the scarce supply of donor livers but requires a tool such as LTCS to predict cirrhosis-related mortality.

Second, the ability to predict longer-term liver-related survival can help clinical management in situations where liver-related mortality may compromise outcomes and/or obviate the benefits of treatment. For example, if a patient with cirrhosis is diagnosed with intermediate-stage lung cancer that is potentially curable by therapy with adverse side effects but the predicted survival is 18 months, the patient may conclude that the risks of aggressive cancer therapy may outweigh the benefits. Conversely, if the predicted 5-year liver-related survival is high, the benefits of aggressive therapy likely outweigh the risks. Or a patient with prostate cancer could consider the expected liver-related survival in the context of treatment decisions that would impact recovery and quality of life (e.g., surgery vs. radiation). As the population with cirrhosis ages and transplantation continues to remain an option for a limited number of patients, such clinical scenarios will become even more common. Although we excluded such patients (e.g., advanced cancer) from our study, we

believe this is a strength because it allowed us to better predict cirrhosis-related survival to help inform such clinical decisions. Had we included such patients, we would not have been able to determine whether their death, especially if limited survival, was from cirrhosis or a preexisting life-limiting comorbidity. Therefore, the LTCS score provides important data that differ from those of other cirrhosis risk scores in that they focus to a greater degree on cirrhosis-related mortality. Third, the ability to predict longer-term survival of patients with cirrhosis could improve the process of advanced care planning, for which deficits are partly attributable to a lack of data beyond short-term outcomes.[13]

The LTCS score can be integrated into routine care. The data elements are collected through regular care, and the laboratory components only require three tests: comprehensive metabolic panel, complete blood count, and prothrombin time/INR. Despite the number of data points, they easily could be entered into an online calculator (in development) and/or integrated into the medical record with direct data uploading (e.g., using a "dot" phrase in Epic with back-end calculation of the LTCS score). Further, our score was developed and validated in populations across the spectrum from academic tertiary care centers to community-based clinics and included an external validation cohort that was nearly 40% female (in contrast to the VHA population).

Our study has limitations. First, the patient population and model of health care in the VHA may not generalize to all populations. However, the short-term survival of patients with decompensated cirrhosis in the VHA mirrors that of other cohorts, and our model performed nearly identically in external validation in a cohort that may be better representative of the broader population with cirrhosis with respect to certain factors (e.g., etiology of liver disease, gender, higher overall survival than the VHA cohort).[19,20] Second, although we excluded patients with major life-limiting medical comorbidities, we included some patients who might have conditions (e.g., diastolic heart failure) that could impact non-liver-related survival but are less reliably ascertained in electronic medical record data. Third, our cancer exclusion was broad; therefore, we may have excluded patients without survival limited by cancer. While this exclusion may impact external validity, it ensured internal validity by excluding patients who could have a non-liver-related condition (i.e., cancer) that could lead to a substantial competing risk of mortality. Fourth, we excluded patients with HCC, though this still allows this score to be applied to the overwhelming majority of patients with cirrhosis. Fifth, the complete-case analysis led to exclusion of many patients. However, this did not lead to biased results (i.e., excellent model performance using multiple imputation) in external validation. Sixth, as is the case with other cirrhosis risk models, we focused on variables at the time of diagnosis and did not account for outcome-modifying treatments (e.g., beta-blockers for clinically significant portal hypertension) or time-varying hepatic decompensation events[12,24] given the use of machine learning models. Lastly, we did not include the location of cirrhosis diagnosis (i.e., inpatient vs. outpatient).

In summary, we developed and validated a risk score with excellent discrimination in predicting intermediate-term and longer-term mortality of patients with cirrhosis. By excluding patients with life-limiting comorbidities, our score can be used to better predict liver-related mortality among LT-eligible patients and broader cohorts with cirrhosis. Future work is needed to determine how incorporating this risk score could impact decisions to refer patients for LT and/or prioritize them in the context of a survival benefit–based approach to allocation.

CONFLICT OF INTEREST

Nothing to report

AUTHOR CONTRIBUTIONS

David Goldberg: Study design, data acquisition, interpretation of data, drafting the article, critical revision, final manuscript approval. Alejandro Mantero: Study design, data acquisition, interpretation of data, drafting the article, critical revision, final manuscript approval. David Kaplan: Study design, data acquisition, interpretation of data, critical revision, final manuscript approval. Cindy Delgado: Study design, data acquisition, critical revision, final manuscript approval. Binu John: Study design, data acquisition, interpretation of data, critical revision, final manuscript approval. Nadine Nuchovich: Study design, data acquisition, critical revision, final manuscript approval. Ezekiel Emanuel: Study design, interpretation of data, drafting the article, critical revision, final manuscript approval. Peter P. Reese: Study design, data acquisition, interpretation of data, drafting the article, critical revision, final manuscript approval.

ETHICS STATEMENTS

The requirement for informed consent was waived for this retrospective study by the IRBs at the University of Miami and the Bruce Carter VA.

DATA AVAILABILITY STATEMENT

The data from the Veterans Health Administration and the OneFlorida Clinical Research Consortium are not publicly available, and the authors are not able to share the data due to restrictions in the data use agreement. However, the source code used to create the final data sets and to run the models are available from the corresponding author, upon reasonable request.

ORCID

Cindy Delgado https://orcid.org/0000-0002-7156-0107
Binu John https://orcid.org/0000-0002-2002-9682

REFERENCES

1. GBD 2017 Cirrhosis Collaborators. The global, regional, and national burden of cirrhosis by cause in 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet Gastroenterol Hepatol. 2020;5(3):245–66.
2. Tapper EB, Parikh ND. Mortality due to cirrhosis and liver cancer in the United States, 1999–2016: observational study. BMJ. 2018;362:k2817.
3. Younossi ZM. Non-alcoholic fatty liver disease—a global public health perspective. J Hepatol. 2019;70(3):531–44.
4. Orman ES, Roberts A, Ghabril M, Nephew L, Desai AP, Patidar K, et al. Trends in characteristics, mortality, and other outcomes of patients with newly diagnosed cirrhosis. JAMA Netw Open. 2019;2(6):e196412.
5. Malinchoc M, Kamath PS, Gordon FD, Peine CJ, Rank J, ter Borg PC. A model to predict poor survival in patients undergoing transjugular intrahepatic portosystemic shunts. Hepatology. 2000;31(4):864–71.
6. Kim WR, Mannalithara A, Heimbach JK, Kamath PS, Asrani SK, Biggins SW, et al. MELD 3.0: the Model for End-Stage Liver Disease updated for the modern era. Gastroenterology. 2021;161(6):1887–1895.e4.
7. Biggins SW, Kim WR, Terrault NA, Saab S, Balan V, Schiano T, et al. Evidence-based incorporation of serum sodium concentration into MELD. Gastroenterology. 2006;130(6):1652–60.
8. Atiemo K, Skaro A, Maddur H, Zhao L, Montag S, VanWagner L, et al. Mortality risk factors among patients with cirrhosis and a low Model for End-Stage Liver Disease sodium score (≤15): an analysis of liver transplant allocation policy using aggregated electronic health record data. Am J Transplant. 2017;17(9):2410–9.
9. Mazumder NR, Atiemo K, Daud A, Kho A, Abecassis M, Levitsky J, et al. Patients with persistently low MELD-Na scores continue to be at risk of liver-related death. Transplantation. 2020;104(7):1413–8.
10. Jepsen P, Vilstrup H, Lash TL. Development and validation of a comorbidity scoring system for patients with cirrhosis. Gastroenterology. 2014;146(1):147–56; quiz e15–6.
11. Kanwal F, Taylor TJ, Kramer JR, Cao Y, Smith D, Gifford AL, et al. Development, validation, and evaluation of a simple machine learning model to predict cirrhosis mortality. JAMA Netw Open. 2020;3(11):e2023780.
12. Trebicka J, Gu W, de Ledinghen V, Aube C, Krag A, Praktiknjo M, et al. Two-dimensional shear wave elastography predicts survival in advanced chronic liver disease. Gut. 2022;71(2):402–14.
13. Patel AA, Ryan GW, Tisnado D, Chuang E, Walling AM, Saab S, et al. Deficits in advance care planning for patients with decompensated cirrhosis at liver transplant centers. JAMA Intern Med. 2021;181(5):652–60.
14. Goldberg D, French B, Newcomb C, Liu Q, Sahota G, Wallace AE, et al. Patients with hepatocellular carcinoma have highest rates of wait-listing for liver transplantation among patients with end-stage liver disease. Clin Gastroenterol Hepatol. 2016;14(11):1638–46.e2.
15. Goldberg DS, French B, Sahota G, Wallace AE, Lewis JD, Halpern SD. Use of population-based data to demonstrate how waitlist-based metrics overestimate geographic disparities in access to liver transplant care. Am J Transplant. 2016;16(10):2903–11.
16. Goldberg DS, Taddei TH, Serper M, Mehta R, Dieperink E, Aytaman A, et al. Identifying barriers to hepatocellular carcinoma surveillance in a national sample of patients with cirrhosis. Hepatology. 2017;65(3):864–74.
17. Mahmud N, Fricker Z, Hubbard RA, Ioannou GN, Lewis JD, Taddei TH, et al. Risk prediction models for postoperative mortality in patients with cirrhosis. Hepatology. 2021;73(1):204–18.
18. Mahmud N, Kaplan DE, Goldberg DS, Taddei TH, Serper M. Changes in hepatocellular carcinoma surveillance and risk factors for noncompletion in the Veterans Health Administration cohort during the coronavirus disease 2019 pandemic. Gastroenterology. 2021;160(6):2162–4.e3.
19. Mahmud N, Kaplan DE, Taddei TH, Goldberg DS. Incidence and mortality of acute-on-chronic liver failure using two definitions in patients with compensated cirrhosis. Hepatology. 2019;69(5):2150–63.
20. Mahmud N, Sundaram V, Kaplan DE, Taddei TH, Goldberg DS. Grade 1 acute on chronic liver failure is a predictor for subsequent grade 3 failure. Hepatology. 2020;72(1):230–9.
21. Serper M, Taddei TH, Mehta R, D'Addeo K, Dai F, Aytaman A, et al. Association of provider specialty and multidisciplinary care with hepatocellular carcinoma treatment and mortality. Gastroenterology. 2017;152(8):1954–64.
22. Goldberg DS, French B, Forde KA, Groeneveld PW, Bittermann T, Backus L, et al. Association of distance from a transplant center with access to waitlist placement, receipt of liver transplantation, and survival among US veterans. JAMA. 2014;311(12):1234–43.
23. Pullen LC. Transplantation remains daunting for many veterans. Am J Transplant. 2017;17(1):1–2.
24. Kanwal F, Hernaez R, Liu Y, Taylor TJ, Rana A, Kramer JR, et al. Factors associated with access to and receipt of liver transplantation in veterans with end-stage liver disease. JAMA Intern Med. 2021;181(7):949–59.
25. Su F, Yu L, Berry K, Liou IW, Landis CS, Rayhill SC, et al. Aging of liver transplant registrants and recipients: trends and impact on waitlist outcomes, post-transplantation outcomes, and transplant-related survival benefit. Gastroenterology. 2016;150(2):441–53.
26. Berry K, Ioannou GN. Comparison of liver transplant–related survival benefit in patients with versus without hepatocellular carcinoma in the United States. Gastroenterology. 2015;149(3):669–80.
27. Su F, Yu L, Berry K, Liou IW, Landis CS, Rayhill SC, et al. Aging of liver transplant registrants and recipients: trends and impact on waitlist outcomes, post-transplantation outcomes, and transplant-related survival benefit. Gastroenterology. 2016;150(2):441–53.e6; quiz e16.
28. 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(3):1144–65.
29. Murphy SP, Ibrahim NE, Januzzi JL Jr. Heart failure with reduced ejection fraction: a review. JAMA. 2020;324(5):488–504.
30. Ahmed A, Aronow WS, Fleg JL. Higher New York Heart Association classes and increased mortality and hospitalization in patients with heart failure and preserved left ventricular function. Am Heart J. 2006;151(2):444–50.
31. Goldberg D, Mantero A, Newcomb C, Delgado C, Forde K, Kaplan D, et al. Development and validation of a model to predict long-term survival after liver transplantation. Liver Transpl. 2021;27(6):797–807.
32. Goldberg D, Mantero A, Newcomb C, Delgado C, Forde KA, Kaplan DE, et al. Predicting survival after liver transplantation in patients with hepatocellular carcinoma using the LiTES-HCC score. J Hepatol. 2021;74(6):1398–406.
33. Levey AS, Eckardt K-U, Tsukamoto Y, Levin A, Coresh J, Rossert J, et al. Definition and classification of chronic kidney disease: a position statement from Kidney Disease: Improving Global Outcomes (KDIGO). Kidney Int. 2005;67(6):2089–100.
34. Wong F, Angeli P. New diagnostic criteria and management of acute kidney injury. J Hepatol. 2017;66(4):860–1.
35. Angeli P, Ginès P, Wong F, Bernardi M, Boyer TD, Gerbes A, et al. Diagnosis and management of acute kidney injury in

patients with cirrhosis: revised consensus recommendations of the International Club of Ascites. J Hepatol. 2015;62(4):968–74.
36. Angeli P, Gines P, Wong F, Bernardi M, Boyer TD, Gerbes A, et al. Diagnosis and management of acute kidney injury in patients with cirrhosis: revised consensus recommendations of the International Club of Ascites. Gut. 2015;64(4):531–7.
37. Penston J. Should we use total mortality rather than cancer specific mortality to judge cancer screening programmes? Yes. BMJ. 2011;343:d6395.
38. Wingfield LR, Ceresa C, Thorogood S, Fleuriot J, Knight S. Using artificial intelligence for predicting survival of individual grafts in liver transplantation: a systematic review. Liver Transpl. 2020;26(7):922–34.
39. Spann A, Yasodhara A, Kang J, Watt K, Wang BO, Goldenberg A, et al. Applying machine learning in liver disease and transplantation: a comprehensive review. Hepatology. 2020;71(3):1093–105.
40. Sanchez-Pinto LN, Stroup EK, Pendergrast T, Pinto N, Luo Y. Derivation and validation of novel phenotypes of multiple organ dysfunction syndrome in critically ill children. JAMA Netw Open. 2020;3(8):e209271.
41. Dugas AF, Hsieh Y-H, LoVecchio F, Moran GJ, Steele MT, Talan DA, et al. Derivation and validation of a clinical decision guideline for influenza testing in 4 US emergency departments. Clin Infect Dis. 2019;70(1):49–58.
42. Rothberg MB, Lindenauer PK, Lahti M, Pekow PS, Selker HP. Risk factor model to predict venous thromboembolism in hospitalized medical patients. J Hosp Med. 2011;6(4):202–9.
43. Ratzinger F, Haslacher H, Perkmann T, Pinzan M, Anner P, Makristathis A, et al. Machine learning for fast identification of bacteraemia in SIRS patients treated on standard care wards: a cohort study. Sci Rep. 2018;8(1):12233.
44. Sanchez-Pinto LN, Venable LR, Fahrenbach J, Churpek MM. Comparison of variable selection methods for clinical predictive modeling. Int J Med Inform. 2018;116:10–17.
45. Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Stat. 2001;29(5):1189–232.
46. Friedman JH. Stochastic gradient boosting. Comput Stat Data Anal. 2002;38(4):367–78.
47. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. Ann Appl Stat. 2008;2(3):841–60.
48. Zou H, Hastie T. Regularization and variable selection via the elastic net. J R Statist Soc B. 2005;67(2):301–20.
49. Simon N, Friedman JH, Hastie T, Tibshirani R. Regularization paths for Cox's proportional hazards model via coordinate descent. J Stat Softw. 2011;39(5):1–13.
50. Breiman L. Random forests. 2001 [2021 Oct 14]. Available from: https://www.stat.berkeley.edu/users/breiman/randomforest2001.pdf.
51. Chambless LE, Diao G. Estimation of time-dependent area under the ROC curve for long-term risk prediction. Stat Med. 2006;25(20):3474–86.
52. Clark TG, Altman DG. Developing a prognostic model in the presence of missing data: an ovarian cancer case study. J Clin Epidemiol. 2003;56(1):28–37.
53. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials—a practical guide with flowcharts. BMC Med Res Methodol. 2017;17(1):162.
54. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009;338:b2393.
55. Stekhoven DJ, Bühlmann P. MissForest—non-parametric missing value imputation for mixed-type data. Bioinformatics. 2012;28(1):112–8.
56. van Houwelingen H, Putter H. Dynamic prediction in clinical survival analysis. Boca Raton, FL: CRC Press; 2011.
57. Spiegelhalter DJ. Probabilistic prediction in patient management and clinical trials. Stat Med. 1986;5(5):421–33.
58. Shenkman E, Hurt M, Hogan W, Carrasquillo O, Smith S, Brickman A, et al. OneFlorida Clinical Research Consortium: linking a clinical and translational science institute with a community-based distributive medical education model. Acad Med. 2018;93(3):451–5.
59. McDonough CW, Babcock K, Chucri K, Crawford DC, Bian J, Modave F, et al. Optimizing identification of resistant hypertension: computable phenotype development and validation. Pharmacoepidemiol Drug Saf. 2020;29(11):1393–401.
60. Vadaparampil ST, Fuzzell LN, Rathwell J, Reich RR, Shenkman E, Nelson DR, et al. HCV testing: order and completion rates among baby boomers obtaining care from seven health systems in Florida, 2015–2017. Prev Med. 2021;153:106222.
61. Andaur Navarro CL, Damen JAA, Takada T, Nijman SWJ, Dhiman P, Ma J, et al. Risk of bias in studies on prediction models developed using supervised machine learning techniques: systematic review. BMJ. 2021;375:n2281.
62. Mahmud N, Hubbard RA, Kaplan DE, Taddei TH, Goldberg DS. Risk prediction scores for acute on chronic liver failure development and mortality. Liver Int. 2020;40(5):1159–67.
63. Rosenblatt R, Wahid N, Halazun KJ, Kaplan A, Jesudian A, Lucero C, et al. Black patients have unequal access to listing for liver transplantation in the United States. Hepatology. 2021;74(3):1523–32.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Goldberg D, Mantero A, Kaplan D, Delgado C, John B, Nuchovich N, et al. Accurate long-term prediction of death for patients with cirrhosis. Hepatology. 2022;00:1–12. https://doi.org/10.1002/hep.32457