JHEP|Reports

AI-Cirrhosis-ECG (ACE) score for predicting decompensation and liver outcomes

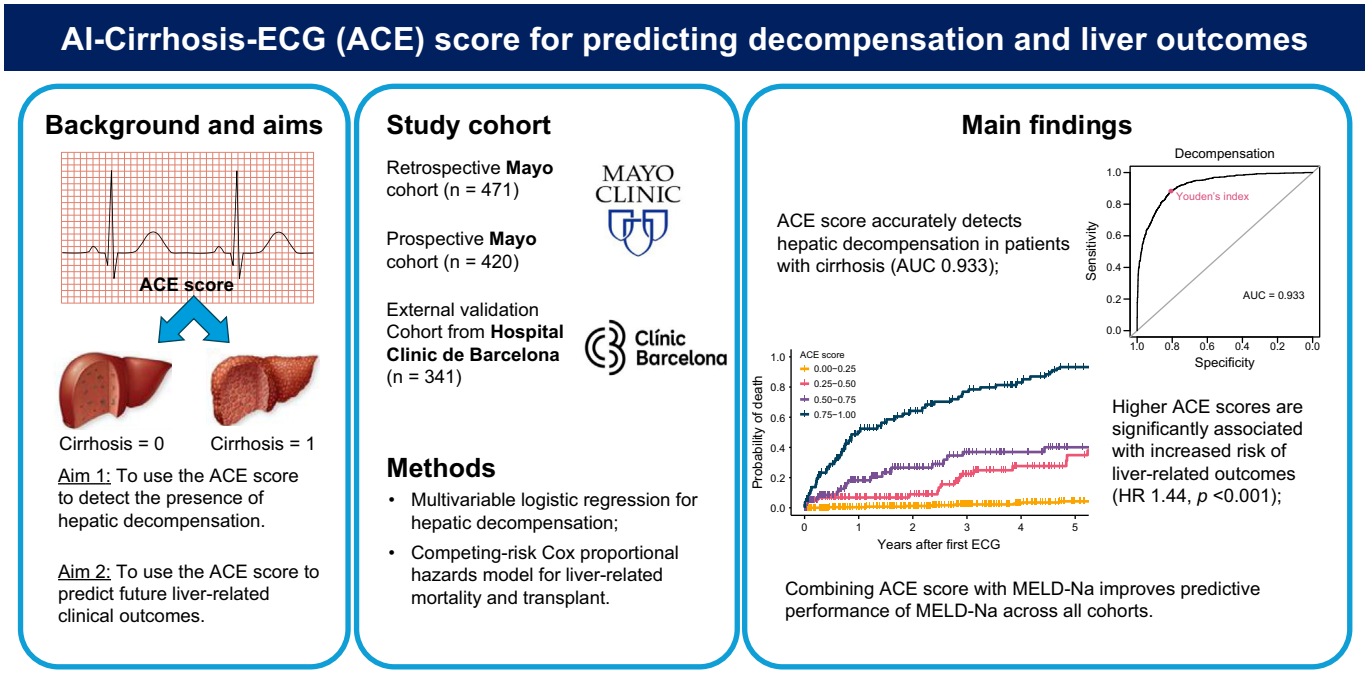
Authors

Joseph C. AhnPuru Rattan, Patrick Starlinger, …, Patrick S. Kamath, Pere Gines, Douglas A. Simonetto

Correspondence

Simonetto.Douglas@mayo.edu (D.A. Simonetto).

Graphical abstract



Highlights:

- ACE score accurately detects hepatic decompensation in patients with cirrhosis (AUC 0.933).
- Higher ACE scores significantly associated with increased risk of liver-related outcomes across multiple cohorts.
- Combining ACE score with MELD-Na improves prediction of adverse clinical outcomes compared to MELD-Na alone.

Impact and implications:

This study demonstrates the potential of artificial intelligence to enhance prognostication in liver disease, addressing the critical need for improved risk stratification in cirrhosis management. The AI-Cirrhosis-ECG (ACE) score, derived from widely available ECGs, shows promise as a non-invasive tool for detecting hepatic decompensation and predicting liver-related outcomes, which could significantly impact clinical decision-making and resource allocation in hepatology. These findings are particularly important for hepatologists, transplant surgeons, and patients with cirrhosis, as they offer a novel approach to complement existing prognostic models such as model for end-stage liver disease-sodium. In practical terms, the ACE score could be integrated into routine clinical assessments to provide more accurate risk predictions, potentially improving the timing of interventions, optimizing transplant listing decisions, and ultimately enhancing patient outcomes. However, further validation in diverse populations and integration with other established predictors is necessary before widespread clinical implementation.

https://doi.org/10.1016/j.jhepr.2025.101356

© 2025 The Author(s). Published by Elsevier B.V. on behalf of European Association for the Study of the Liver (EASL). This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

JHEP|Reports

AI-Cirrhosis-ECG (ACE) score for predicting decompensation and liver outcomes

Joseph C. Ahn[1,†], Puru Rattan[1,†], Patrick Starlinger[2], Adrià Juanola[3,4,5], Maria José Moreta[3,4,5], Jordi Colmenero[3,4,5,6], Bashar Aqel[7], Andrew P. Keaveny[8], Aidan F. Mullan[9], Kan Liu[10], Zachi I. Attia[10], Alina M. Allen[1], Paul A. Friedman[10], Vijay H. Shah[1], Peter A. Noseworthy[10], Julie K. Heimbach[11], Patrick S. Kamath[1], Pere Gines[3,4,5,6], Douglas A. Simonetto[1,*]

JHEP Reports **2025**. vol. 7 | 1–10

Check for updates

Background & Aims: Accurate prediction of disease severity and prognosis are challenging in patients with cirrhosis. We evaluated whether the deep learning-based AI-Cirrhosis-ECG (ACE) score could detect hepatic decompensation and predict clinical outcomes in cirrhosis.

Methods: We analyzed 2,166 ECGs from 472 patients in a retrospective Mayo Clinic cohort, 420 patients in a prospective Mayo transplant cohort, and 341 patients in an external validation cohort from Hospital Clínic de Barcelona. The ACE score's performance was assessed using receiver-operating characteristic analysis for decompensation detection and competing risks Cox regression for outcome prediction.

Results: The ACE score showed high accuracy in detecting hepatic decompensation (area under the curve 0.933, 95% CI: 0.923–0.942) with 88.0% sensitivity and 84.3% specificity at an optimal threshold of 0.25. In multivariable analysis, each 0.1-point increase in ACE score was independently associated with increased risk of liver-related death (hazard ratio [HR] 1.44, 95% CI 1.32–1.58, p <0.001). Adding ACE to model for end-stage liver disease-sodium significantly improved prediction of adverse outcomes across all cohorts (c-statistics: retrospective cohort 0.903 *vs.* 0.844; prospective cohort 0.779 *vs.* 0.735; external validation 0.744 *vs.* 0.732; all p <0.001).

Conclusions: The ACE score accurately identifies hepatic decompensation and independently predicts liver-related outcomes in cirrhosis. This non-invasive tool enhances current prognostic models and may improve risk stratification in cirrhosis management.

© 2025 The Author(s). Published by Elsevier B.V. on behalf of European Association for the Study of the Liver (EASL). This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Chronic liver diseases, and their common end-stage manifestations of cirrhosis, impose a massive burden on healthcare worldwide[1,2] that is only predicted to increase in the coming decades.[1] Initially, cirrhosis is a surreptitious process present in a relatively asymptomatic or 'compensated' state, which becomes conspicuous over time with the development of ascites, variceal hemorrhage, or hepatic encephalopathy.[3,4] The presence of any of these complications indicates the transition to 'decompensated cirrhosis', a rapidly progressive state with a marked reduction in survival.[5] This evolution in the disease process is closely linked to portal pressure, with worsening portal hypertension directly associated with the severity of disease.[3,6] For patients with cirrhosis, accurate estimation of both disease severity and their prognosis are of the utmost importance, allowing not only for timely implementation of preventative therapies, but also for appropriate organ allocation among candidates for liver transplantation (LT).[7] However, various iterations of the model for end-stage liver disease (MELD) score with its relationship to survival prognosis[8–10] have become the basis of LT organ allocation over the last two decades. However, even with iterative changes the MELD remains an imperfect predictor of waitlist mortality, spurring numerous recent efforts to improve its performance.[11–14] Discoveries underlying the pathophysiology of portal hypertension over the last 50 years have revealed the profound involvement of the cardiovascular system in cirrhosis, especially the significant roles of vascular tone, in the splanchnic and intrahepatic circulation, and of the hyperdynamic circulatory state present in decompensated cirrhosis.[15,16] Recently, our group developed the AI-Cirrhosis-Electrocardiogram (ACE) score, a deep learning-based artificial intelligence (AI) system utilizing a convolutional neural network (CNN) that is able to discriminate between patients with and without cirrhosis.[17] Formulated using ECGs from 25,904 patients with and without advanced cirrhosis, this algorithm analyzes the raw signal from standard 12-lead ECGs to produce a numeric ACE score between 0 and 1 whose magnitude reflects the presence of cirrhosis-related

* Corresponding author. Address: Division of Gastroenterology and Hepatology, Mayo Clinic Rochester, 200 First Street SW, Rochester, MN 55905, USA. Phone: (507) 284-2511, Fax: (507) 538-7802.
E-mail address: Simonetto.Douglas@mayo.edu (D.A. Simonetto).
† These authors contributed equally to this study.
https://doi.org/10.1016/j.jhepr.2025.101356

ECG changes. While this proof of concept study demonstrated that this algorithm had excellent discriminative performance, intriguingly, secondary analyses of the results revealed an association of the ACE score with severity of liver disease, specifically mirroring the progression and resolution of disease in patients who approached and subsequently underwent LT.[17] In this study, we aim to evaluate the ACE score's ability to (1) detect the presence of hepatic decompensation, and (2) improve prediction of liver-related outcomes in patients with cirrhosis. We investigate these capabilities in three separate cohorts and examine how incorporating the ACE score can enhance the predictive performance of the existing MELD-sodium (MELD-Na) score.

Patients and methods

Patient cohorts and data collection

Retrospective Mayo cohort

In order to analyze the ACE score's relationship with disease stage and longitudinal outcomes, a foundational cohort of patients who initially presented to the Mayo Clinic with compensated cirrhosis was selected. This 'retrospective Mayo cohort' consisted of patients who were seen in the Division of Gastroenterology and Hepatology at Mayo Clinic, Rochester between January 1, 1995 and December 31, 2020 and were documented to have 'compensated cirrhosis' by a hepatologist using an internal clinical documentation search tool. Individuals with cardiac cirrhosis and those who went on to develop hepatocellular carcinoma or cholangiocarcinoma were excluded. Subsequently, any clinical documentation following the initial encounter for compensated cirrhosis, of portal hypertensive ascites, variceal hemorrhage, overt hepatic encephalopathy, hepatorenal syndrome, or spontaneous bacterial peritonitis resulted in a decompensated label. After finalization of the cohort, each patient was classified into the following outcome categories: (1) LT, (2) death, (3) alive or lost to follow-up. The ACE score was applied to all available ECGs from the date of each patient's initial documentation of compensated cirrhosis till date of outcome. Relevant laboratory values within 90 days of each ECG were also extracted from the electronic health record (EHR). Additionally, billed diagnosis codes from the EHR were used to classify the presence of relevant comorbidities for each patient using the Charlson and Elixhauser Comorbidity Index groupings.[18]

Prospective Mayo cohort

Subsequently, to internally validate the association of ACE and liver-related outcomes, we identified an independent cohort of patients not included in the 'retrospective Mayo cohort' with decompensated cirrhosis who were waitlisted for LT at one of the Mayo Clinic transplant centers. These patients were prospectively included during their evaluation visit between January 1, 2016 and December 31, 2019 to be in a cohort to study outcomes in waitlisted LT candidates. All patients in this 'prospective Mayo cohort' had an ECG performed at the time of evaluation and the necessary laboratory blood tests needed to calculate the MELD-Na score. Patients with non-cirrhotic etiologies of liver disease, cardiac cirrhosis, hepatocellular carcinoma, or cholangiocarcinoma were excluded. The ACE model was again applied to any available ECGs performed for this cohort from the time of LT evaluation to the dates of their last follow-up, liver-related death, or LT, producing an ACE score for each ECG.

External validation cohort from Hospital Clínic de Barcelona

For external validation of the ACE score outside of Mayo Clinic, the ACE model was applied to ECGs from a cohort of patients from Hospital Clínic de Barcelona in Barcelona, Catalonia, Spain. Patients were selected from the registry of patients evaluated for LT between 2013 and 2022 in Hospital Clínic de Barcelona. Patients were screened during hospitalization for an acute decompensation episode or during routine outpatient clinic visits. Those with decompensated cirrhosis and available ECGs were included in the study, irrespective of their LT listing status. Patients with acute liver failure, previous LT or hepatocellular carcinoma were excluded from the analysis. The current study was approved by the local ethics committee.

Data analysis

Association between the ACE score and hepatic decompensation at the time of ECG was assessed both by comparing the distributions of ACE in these subsets, and using univariable and multivariable logistic regression adjusting for MELD-Na and covariates including demographics, comorbidities, and etiology of cirrhosis. The distributions of ACE scores were compared using the Kruskal–Wallis test followed by *post-hoc* pairwise comparisons between each combination. In the logistic regression analyses, to account for repeated ECG measures from the same patient, a random intercept was included in the model for each unique patient while a least absolute shrinkage and selection operator (LASSO) penalty was applied to multivariable regression to minimize overfitting of the data. The ability of the ACE score to identify decompensation was evaluated using a receiver-operating characteristic (ROC) curve and the resulting area under the ROC curve (AUC). A confidence interval for the AUC was calculated using a bootstrap resampling estimator. The optimal ACE cut-off for both binary classification models was determined using Youden's Index to provide a balance between sensitivity and specificity. The final threshold was rounded to the nearest 0.05 for ease of identification.

The effect of the ACE score on liver-related outcomes, including liver-related death and LT, was evaluated with Kaplan–Meier survival analyses and cause-specific Cox proportional hazards models. For the Kaplan–Meier analyses, patients were divided into four quartiles based on their baseline ACE score (0.00–0.25, 0.25–0.50, 0.50–0.75, and 0.75–1.00) and survival was compared using cumulative incidence curves for both outcomes. In the Cox models, the ACE score was treated as a time-dependent covariate, updated when new ECG data became available. Other clinical parameters, including MELD-Na and its component laboratory values, were also updated using the values closest to the time of each ECG, provided they were available within 3 months of the ECG date. This approach ensures that our analysis reflects the most current clinical status of each patient at the time of ECG measurement.

The primary outcomes of interest were LT and liver-related death. All patients were censored at the time of last follow-up,

death from any cause, and LT. The Cox regression models used were univariable, multivariable adjusting for MELD-Na, and multivariable adjusting for patient demographics, comorbidities, liver etiology, and MELD-Na. The associations between ACE score and liver-related outcomes in patients deemed to be low-risk by MELD-Na were assessed with a sensitivity analysis where the cohort was restricted to only patients from the prospective cohort with MELD-Na ≤15.

In all the above analyses, a value of p <0.05 was considered statistically significant. Statistical modeling and analyses were conducted using R 4.0.3 (R Foundation for Statistical Computing, Vienna, Austria). All the EHR data extracted and analyzed for this study were deidentified, and thus the study was deemed exempt by the institutional review board at the Mayo Clinic.

Model explainability via saliency maps

To provide insight into the features learned by the deep learning model, we generated gradient-based saliency maps for selected ECG tracings. These saliency maps highlight the portions of the input signal to which the model is most 'attentive' when predicting cirrhosis-related risk. Specifically, we used a gradient-based visualization method (*e.g.* Grad-CAM or an integrated gradients approach) applied at the final convolutional layers of the network. In this way, the algorithm's focus on particular segments of the ECG waveform (*e.g.* QRS complexes, T waves) became more apparent, offering a degree of interpretability regarding how the ACE score is derived.

Results

Retrospective Mayo cohort

Our search strategy initially identified 572 patients with documented compensated cirrhosis, after exclusions this resulted in a total of 472 patients comprising our main retrospective Mayo cohort of patients (Table 1). This cohort had 2,166 ECGs available for application of the ACE model. There were slightly fewer females than males (42.8% *vs.* 57.2%) and the cohort was predominantly white (89.6%). Metabolic dysfunction-associated steatotic liver disease (MASLD) (37.9%) was the most prevalent etiology of liver disease among these individuals. The median age at the time of ECG was 64 years.

Cross-sectional association of ACE score with decompensation (retrospective Mayo cohort)

To explore the potential application of the ACE score in identifying patients at risk of decompensation, we analyzed the retrospective Mayo cohort, which included patients who transitioned from compensated to decompensated cirrhosis. Among the 472 patients in the retrospective cohort, 190 (40.3%) had ECGs performed in a decompensated state. Conversely, 282 (59.7%) patients had ECGs only available from periods when they were compensated, while 59 (12.5%) had ECGs representing both a compensated and decompensated stage of their disease. Of the 2,166 ECGs available for the retrospective cohort, 1,243 represented a compensated state and 923 occurred after a decompensating event. The median ACE scores for ECGs from compensated, and decompensated patients were 0.0344, and 0.671 respectively (Fig. 1A). The distribution of ACE scores within these two stages of disease were overall statistically different by a Kruskal–Wallis test (p <0.01), whereas subsequent pairwise tests also showed significant difference in ACE between each group (p <0.0001 for all pairs).

Evaluation of the ACE score's ability to detect hepatic decompensation revealed excellent discriminative performance, with an AUC of 0.933 (95% CI: 0.923–0.942) (Fig. 1B). The optimal threshold for decompensated disease was an ACE score of 0.25, which yielded a sensitivity of 88.0% (95% CI 85.7–89.9%) and a specificity of 84.3% (95% CI 82.1–86.3%).

Table 1. Baseline patient characteristics in Mayo Clinic cohorts.

	Retrospective Mayo cohort (n = 472)	Prospective Mayo cohort (n = 421)
Median Age at ECG (Q1, Q3)	64 (59, 72)	58 (50, 65)
Sex, n (%)		
Male	270 (57.2)	236 (56.1)
Female	202 (42.8)	185 (43.9)
Race, n (%)		
American Indian/Alaskan Native	4 (0.8)	11 (2.6)
Asian	9 (1.9)	9 (2.1)
Black	11 (2.3)	8 (1.9)
White	423 (89.4)	353 (83.8)
Other/did not disclose	25 (5.3)	40 (9.5)
Comorbidities, n (%)		
Coronary artery disease	55 (11.6)	31 (7.4)
Hypertension	322 (68.1)	213 (50.6)
Diabetes mellitus	208 (44.0)	150 (35.6)
Chronic kidney disease	121 (25.6)	188 (44.7)
Chronic lung disease	197 (41.6)	79 (18.8)
Etiology of liver disease, n (%)		
Viral	138 (29.2)	50 (11.9)
Alcohol-related	110 (23.3)	113 (26.8)
Non-alcoholic steatohepatitis	179 (37.8)	122 (29.0)
Autoimmune	30 (6.3)	15 (3.6)
Biliary	33 (7.0)	62 (14.7)
Genetic/metabolic	23 (4.9)	10 (2.4)
Cryptogenic	41 (8.7)	47 (11.2)

Descriptive statistics, no statistical tests applied.

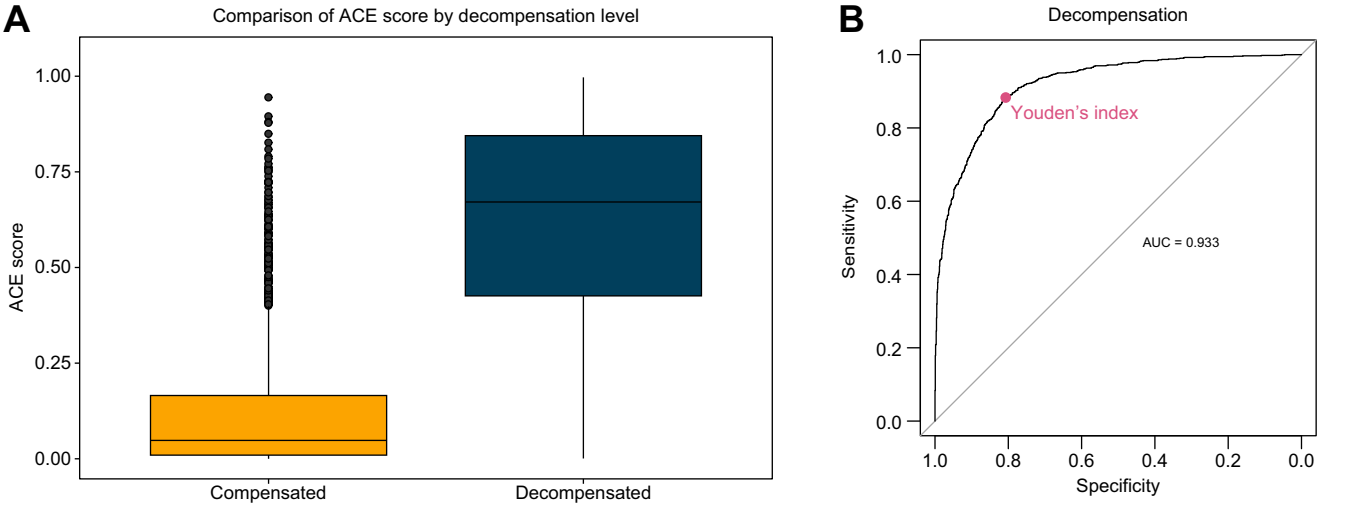


Fig. 1. The ACE score's association with hepatic decompensation on 12-lead ECGs. (A) Distribution of the ACE score among compensated and decompensated patients. Statistical significance determined using the Kruskal–Wallis test (p <0.01) followed by pairwise comparisons (p <0.0001 for all pairs). (B) Receiver operator characteristic curve for the ACE score's performance in discriminating compensated and decompensated patients; optimal threshold = 0.25. Area under the curve (AUC) = 0.933 (95% CI 0.923–0.942). ACE, AI-Cirrhosis-ECG; AUC, area under the curve; ECG, electrocardiogram.

Furthermore, using a univariate logistic regression, a 0.1-point increase in ACE score was associated with 4.3 times greater odds of the ECG being from a patient with decompensated disease (Table S1; odds ratio [OR] = 4.23, 95% CI 3.46–5.17, p <0.001). These odds remained significant after adjusting the model for MELD-Na (OR = 4.78, 95% CI 3.59–6.36, p <0.001). A multivariable model fully adjusting for patient demographics, comorbidities, liver disease etiology, and MELD-Na showed that a 0.1-point increase in ACE score was still associated with almost twice as high odds of the ECG belonging to a patient with decompensated disease (OR = 1.91, 95% CI 1.43–2.55, p <0.001) (Table S1).

Longitudinal prediction of outcomes: retrospective Mayo cohort

Building on the ACE score's ability to detect decompensation, we further investigated its potential to predict liver-related outcomes in this cohort, which represents patients at various stages of cirrhosis progression. In the retrospective cohort of 472 patients, a total of 112 (23.7%) patients died as a result of liver-related causes and 42 (8.9%) patients received LT. Significant increases in liver-related mortality were observed across increasing ACE quartiles over 5 years of follow-up (Fig. 2A). Notably, patients in the lowest ACE quartile (0.00–0.25) showed excellent prognosis, with over 90% survival at 5 years post-ECG. In stark contrast, patients in the highest quartile (≥0.75) exhibited poor outcomes, with approximately 50% mortality at 1 year and over 90% mortality 5 years post-ECG (p <0.001). However, with an overall low rate of LT in this cohort, no clear differences in the risk of LT among the ACE quartiles were apparent on the Kaplan–Meier analysis (Fig. 2B).

When the outcomes were modeled using competing risks Cox regression (Table 2), an unadjusted, univariable model found that a 0.1-point increase in ACE score was associated with 59% increased risk of liver-related death (hazard ratio [HR] = 1.59, 95% CI 1.48–1.70, p <0.001) and 40% increased risk of LT (HR = 1.40, 95% CI 1.20–1.63, p <0.001). After adjusting for MELD-Na, a 0.1-point increase in ACE score was still significantly associated with 40% increased risk of liver-related death (HR = 1.40, 95% CI 1.29–1.52, p <0.001), and a marginally increased risk of LT (HR = 1.19, 95% CI 1.01–1.40, p = 0.038). In the full multivariable model, a 0.1-point increase in ACE score was associated with 44% increase in the risk of liver-related death (HR = 1.44, 95% CI 1.32–1.58, p <0.001) but only a non-significant trend towards an increased likelihood of LT (HR = 1.18, 95% CI 0.99–1.40, p = 0.069).

Additionally, comparing the contribution of ACE and MELD-Na to a competing risks model's performance using likelihood ratio tests demonstrated that although ACE and MELD-Na alone had equivalent performance, the concordance of a model containing both was far better (Table 3). Specifically, the performance of a model containing MELD-Na to predict liver-related death in the presence of a competing risk for LT was significantly improved with the addition of the ACE score (C-statistic = 0.844 *vs.* 0.903, p <0.0001). This improvement was also present when evaluating models predicting liver transplant with a competing risk for liver-related death (C-statistic = 0.844 *vs.* 0.861, p = 0.036).

Internal validation: prospective Mayo cohort

To evaluate the ACE score's capability to enhance outcome prediction for waitlisted patients, we analyzed a prospective cohort of patients evaluated for LT. This cohort allows us to assess how incorporating the ACE score might improve upon the predictive performance of MELD-Na in risk stratification for this high-risk population. The prospective Mayo cohort comprised 420 patients recruited at the time of their LT evaluation for decompensated cirrhosis (Table 1). With a median age of 58 years and a slight male predominance (56.1% *vs.* 43.9% female), this cohort represented a diverse array of liver disease etiologies, with alcohol-related (26.8%) and MASLD (29.0%) being the most prevalent, followed by biliary, viral, cryptogenic, and other etiologies.

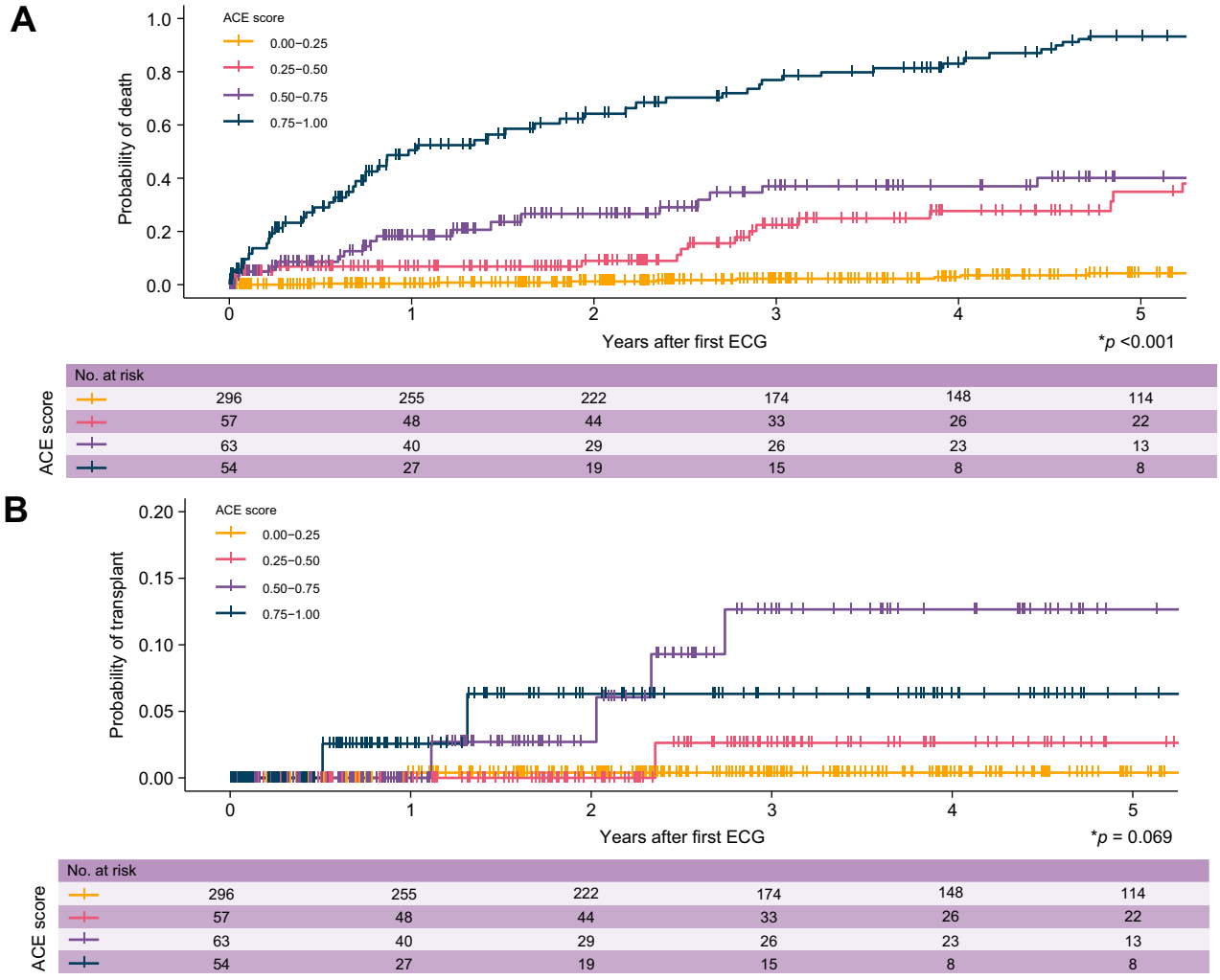


Fig. 2. Kaplan–Meier survival analyses by quartiles of ACE score in the retrospective Mayo cohort. (A) Probability of liver-related death over time. Statistical significance determined using the log-rank test (p <0.001). (B) Probability of liver transplant over time. Statistical significance determined using the log-rank test (p = 0.069). ACE, AI-Cirrhosis-ECG; ECG, electrocardiogram.

Longitudinal outcomes in the prospective Mayo cohort

In the prospective cohort of 420 patients evaluated for LT, a total of 109 (26.0%) patients experienced death from liver-related causes and 221 (52.6%) patients underwent successful LT. On Kaplan–Meier analysis, patients in higher ACE quartiles had significantly higher probabilities of death and LT compared with patients with lower ACE quartiles (Fig. 3A and B).

Table 2. Cox regression for association between ACE score and liver-related outcomes.

	Liver-related death			Liver transplant		
Model	**HR***	**95% CI**	***p* value**	**HR***	**95% CI**	***p* value**
Retrospective Mayo cohort						
Unadjusted	1.59	1.48–1.70	<0.001	1.40	1.20–1.63	<0.001
Adjusted for MELD-Na	1.40	1.29–1.52	<0.001	1.19	1.01–1.40	0.038
Fully adjusted†	1.44	1.32–1.58	<0.001	1.18	0.99–1.40	0.069
Prospective Mayo cohort						
Unadjusted	1.31	1.21–1.41	<0.001	1.31	1.23–1.38	<0.001
Adjusted for MELD-Na	1.21	1.12–1.31	<0.001	1.23	1.17–1.32	<0.001
Fully adjusted†	1.19	1.10–1.29	<0.001	1.25	1.18–1.33	<0.001

Cox proportional hazards models were used. Hazard ratios (HRs) with 95% CIs and p values are provided for each model.
ACE, AI-Cirrhosis-ECG; HR, hazard ratio; MELD-Na, model for end-stage liver disease-sodium score.
*For each 0.1-point increase in ACE score.
†Adjusted for age, sex, race, medical comorbidities, etiology of liver disease, and MELD-Na.

In this prospective cohort, modeling the outcomes with Cox competing risks regression (Table 2), a 0.1-point increase in ACE score was significantly associated with increased risk of liver-related death in the unadjusted univariable model (HR = 1.31, 95% CI 1.22–1.84, p <0.001), the model adjusted for MELD-Na (HR = 1.21, 95% CI 1.11–1.31, p <0.001), and the fully adjusted multivariable model (HR =1.19, 95% CI 1.09–1.29, p <0.001). Furthermore, a 0.1-point increase in ACE score was also associated with an increased probability of LT in the unadjusted univariable model (HR = 1.31, 95% CI 1.22–1.53, p <0.001), the model adjusted for MELD-Na (HR = 1.23, 95% CI 1.17–1.32, p <0.001), and the fully adjusted multivariable model (HR = 1.25, 95% CI 1.18–1.33, p <0.001).

The contribution of ACE alone was lower when compared to MELD-Na alone, to the overall and specific liver-related prediction in the competing risks survival models (Table 3). However, models containing both MELD-Na and ACE performed significantly better than MELD-Na alone in predicting liver-related death (C-statistic = 0.818 *vs.* 0.794, p <0.0001) and liver transplant (C-statistic = 0.764 *vs.* 0.713, p <0.0001).

Notably, when a subset of the prospective cohort with an initial MELD-Na of ≤15 (n = 127) was used to model the association between ACE scores and liver-related outcomes, the contribution of ACE alone was superior to that of MELD-Na (C-statistic = 0.749 *vs.* 0.730, p <0.001) (Table 3). Moreover, addition of ACE to MELD-Na led to significant improvement in prediction of liver-related outcomes compared with MELD-Na alone (liver-related death C-statistic = 0.871 *vs.* 0.778, p <0.001; liver transplant C-statistic = 0.793 *vs.* 0.717, p <0.001).

External validation in the Hospital Clínic de Barcelona cohort

To further validate the ACE score's predictive capabilities in waitlisted patients and assess its generalizability across different populations, we analyzed an external cohort from Hospital Clínic de Barcelona. A total of 341 patients were included in the external validation cohort from Hospital Clínic de Barcelona (Table S2). Median age at the time of ECG was 58 (IQR 52–63) and 73.3% were male. The most common etiologies were alcohol-related (208 patients, 61.0%), viral (41 patients, 12.0%), and MASLD (40 patients, 11.7%). Median MELD-Na was 17 (IQR 13–23). All of them had at least one previous decompensation episode, with ascites being the most common type of decompensation (227, 65%), followed by hepatic encephalopathy (165, 47%) and bacterial infection (117, 34%).

On Kaplan–Meier analysis, similar to the findings in both Mayo cohorts, patients in higher ACE quartiles had significantly higher probabilities of adverse liver-related outcomes compared with patients with lower ACE quartiles (Fig. 4). Unadjusted and adjusted Cox regression models found that ACE was a significant predictor of adverse liver outcomes (Table S3). A 0.1-point increase in ACE was associated with a 9.5% increase in risk for liver-related mortality or transplant (HR = 1.095, 95% CI 1.056–1.134, p <0.001). After adjusting for MELD-Na and age, sex, and liver disease etiology, a 0.1-point increase in ACE was still associated with a 6.4% increase in the risk for adverse liver-related outcomes (HR = 1.064, 95% CI 1.022–1.108, p <0.001). A likelihood ratio test found that the combined Cox model with ACE and MELD-Na significantly outperformed the univariable models with either ACE or MELD-Na (likelihood ratio test [LRT] p <0.001 for both) in this cohort (Table S4).

Explainability: saliency map findings

To further understand the ACE model's decision-making process, we generated saliency maps from representative ECG tracings in each cohort. These maps consistently showed pronounced highlights in regions corresponding to the QRS complex and, in some instances, the T wave (Figs S1–S4). Notably, the model appeared to focus on the morphological variations in the R wave (and occasionally the S wave) across the precordial leads (V3–V6). Similar patterns were observed in the limb leads, although to a varying extent. These patterns suggest that the network is leveraging subtle conduction or repolarization changes that may be more prominent in advanced liver disease.

Across multiple beats, the model's emphasis on these ECG segments remained consistent, reinforcing the idea that cirrhosis-related physiologic changes may manifest as predictable alterations in the QRS complex and repolarization patterns. Although the precise physiologic underpinnings warrant further investigation, the saliency maps imply that the ACE score partly captures the interplay of cardiac structural or autonomic changes associated with cirrhosis.

Table 3. Competing risks Cox models performance in predicting liver-related outcomes.

	Concordance			
Model	**Overall**	**Liver-related death**	**Liver transplant**	***p* value**
Retrospective Mayo cohort				
ACE alone	0.871	0.879	0.799	<0.001
MELD-Na alone	0.844	0.844	0.844	<0.001
ACE + MELD-Na	0.903	0.908	0.861	<0.001
Prospective Mayo cohort				
ACE alone	0.693	0.680	0.698	<0.001
MELD-Na alone	0.735	0.794	0.713	<0.001
ACE + MELD-Na	0.779	0.818	0.764	<0.001
Prospective Mayo cohort with MELD <15				
ACE alone	0.749	0.786	0.739	<0.001
MELD-Na alone	0.730	0.778	0.717	<0.001
ACE + MELD-Na	0.809	0.871	0.793	<0.001

Competing risks Cox models were used. Concordance (C-statistic) values are provided for each model. Statistical significance of model comparisons determined using likelihood ratio tests (*p* values provided).
ACE, AI-Cirrhosis-ECG; MELD-Na, model for end-stage liver disease-sodium score.

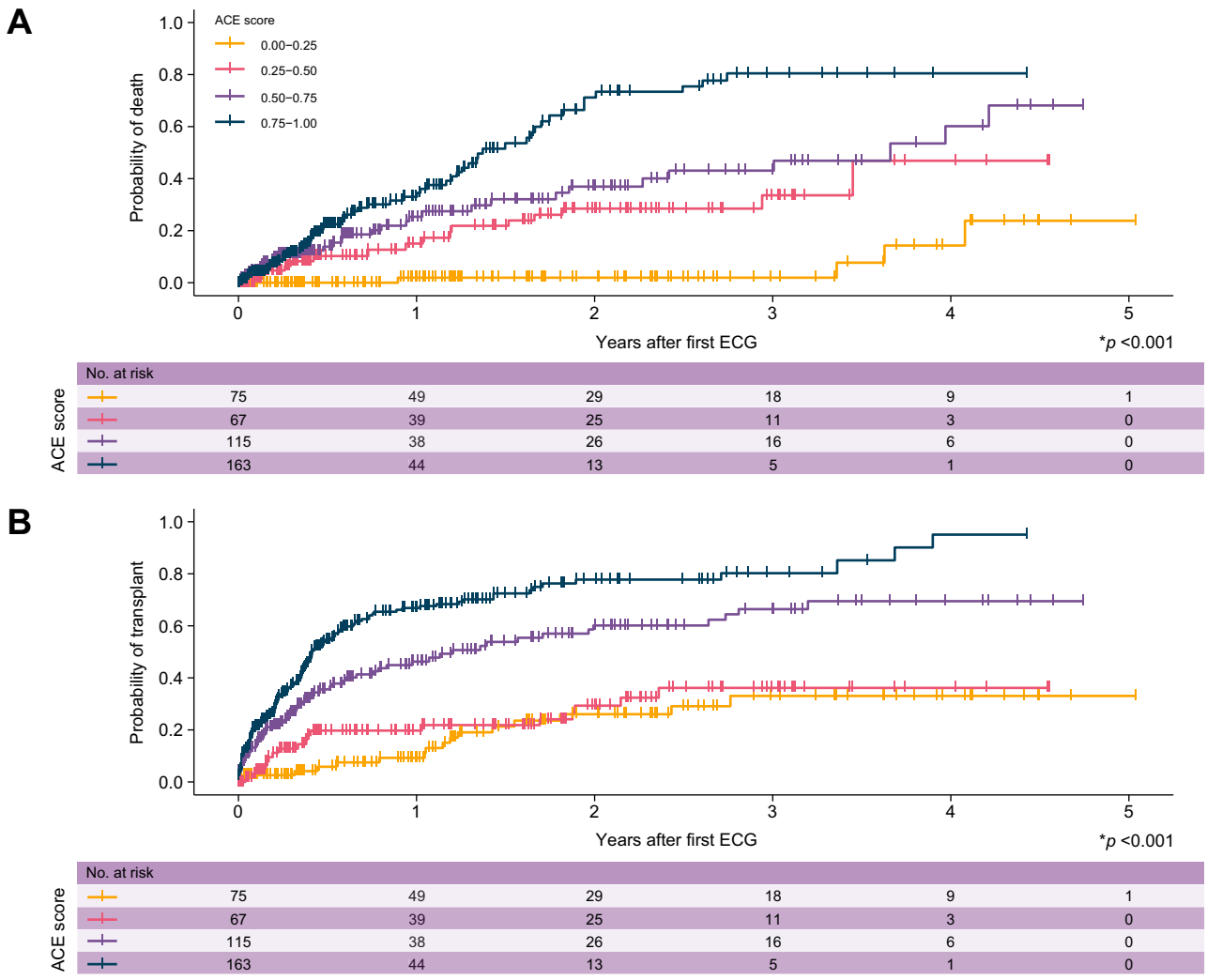


Fig. 3. Kaplan–Meier survival analyses by quartiles of ACE score in the prospective Mayo cohort. (A) Probability of liver-related death over time. Statistical significance determined using the log-rank test (p <0.001). (B) Probability of liver transplant over time. Statistical significance determined using the log-rank test (p <0.001). ACE, AI-Cirrhosis-ECG; ECG, electrocardiogram.

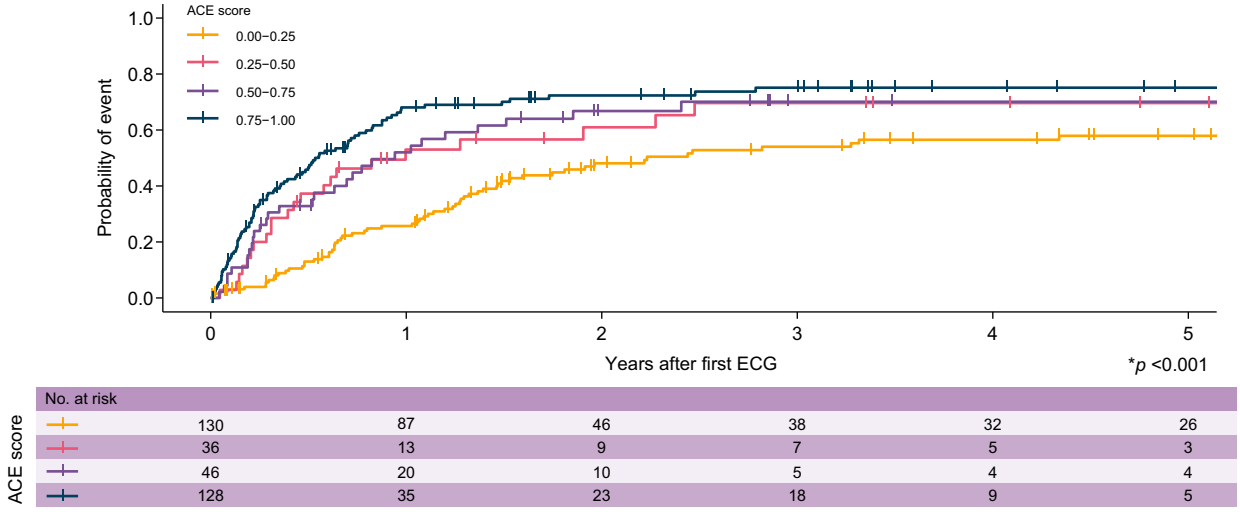


Fig. 4. Kaplan–Meier survival analyses by quartiles of ACE score in the Barcelona external validation cohort. Probability of liver-related death or transplant over time. Statistical significance determined using the log-rank test (p <0.001). ECG, electrocardiogram.

Discussion

In this study, the ACE score,[17] a deep learning-based AI model, demonstrated strong predictive capability for hepatic decompensation and liver-related outcomes, showing potential to enhance current prognostic tools in cirrhosis management. This result was demonstrated in three separate cohorts of patients with cirrhosis: retrospective Mayo cohort, prospective Mayo cohort, and external validation in a cohort of patients from Hospital Clínic de Barcelona, Catalonia, Spain. Additionally, the ACE score has shown the potential to augment existing scores, such as the MELD, to improve prognostication in patients with cirrhosis. Although the exact subtleties in the raw ECG signals that compose the ACE score are not overtly apparent, the significant associations between ACE score and hepatic decompensation provide further evidence that the ACE score derives from various physiologic changes that take place in cirrhosis. While our study provides evidence of the ACE score's predictive capabilities, it also lays the groundwork for future research. The next steps should focus on developing and validating comprehensive predictive models that incorporate the ACE score alongside other established predictors of adverse outcomes in cirrhosis.

This study highlights the potential for AI to improve the clinical assessment and management of patients with cirrhosis. Recent breakthroughs in the field of AI fueled by state-of-the-art machine learning algorithms including deep learning have enabled computers to process complex, high-dimensional data with unprecedented ease. AI is being rapidly adopted in almost all fields of medicine, with hepatology being no exception.[19] With an array of multimodal data applicable to hepatology, there has been an increased focus on the development of portal hypertension and outcome-related models by leveraging a combination of EHR, radiological, and histopathological data.[20–22] To date, this is the first study to apply an AI model for assessing the risk of hepatic decompensation and liver-related outcomes using ECGs. The feasibility of applying deep learning-based AI models for automated interpretation of ECGs is supported by recently published algorithms for prediction of various cardiac and non-cardiac conditions including left ventricular dysfunction,[23] atrial fibrillation,[24] hypertrophic cardiomyopathy,[25] hyperkalemia,[26] as well as sex and age[27] on ECGs.

The results of our study suggest that calculating the ACE score from an ECG in a patient with chronic liver disease may provide a powerful, non-invasive tool for disease staging and risk stratification. The ACE score demonstrated excellent discriminative ability in detecting hepatic decompensation, with an AUC of 0.933 and an optimal decompensation threshold of 0.25. These performance characteristics suggest that the ACE score could potentially serve as an objective, quantitative measure to complement clinical assessment of decompensation. Although the clinical detection of decompensated disease is often apparent, the ability to objectively quantify and potentially predict this transition using a widely available, non-invasive test such as an ECG is particularly significant. The ACE score's performance in this regard suggests its potential to enhance current prognostic tools and improve risk stratification in cirrhosis management.

It is important to note that although our study demonstrates strong predictive potential of the ACE score, it represents an initial step towards developing a comprehensive predictive model. Further research is needed to validate these findings in diverse populations and to integrate the ACE score with other established predictors of adverse outcomes in cirrhosis. This future work will be crucial in translating the promising results of our study into a validated predictive model suitable for clinical implementation. Currently, the MELD score is used to rank those awaiting allocation of donor livers.[13] The increasing size of the waitlist at transplant centers has resulted in numerous efforts to adjunctive models to improve allocation and reduce waitlist mortality. Although the objective laboratory values in the MELD score are easily verifiable for the purpose of allocation policy, several subjective cirrhosis and portal hypertension-related complications, such as degree of hepatic encephalopathy, refractory ascites and even sarcopenia with frailty[11] are known to be highly associated with worse prognosis in advanced liver disease. Since the transition from the Child-Turcotte-Pugh score to the MELD for organ allocation in early 2000s, these important yet subjective prognostic factors have not been captured when prioritizing transplant recipients. In this study, we were able to show that the ACE score is significantly associated with liver-related mortality or the need for liver transplant, independent of MELD. Moreover, the performance of MELD-based models is significantly improved with the addition of ACE. This is well illustrated on a subgroup analysis of patients with MELD-Na less than or equal to 15 who experienced short-term mortality in our cohort and would have been appropriately identified as 'high-risk' for death based on their ACE scores. The strong correlation of ACE with hepatic decompensation and outcomes suggests that the ACE score is perhaps able to capture the subjective clinical complications impacting survival in cirrhosis. The next phase of research should focus on integrating the ACE score with other established predictors of decompensation and adverse outcomes in cirrhosis, such as albumin, bilirubin, platelet count, presence of esophageal varices, and liver stiffness measurements. By combining the ACE score with these traditional markers, we aim to develop and validate enhanced predictive models that could significantly improve risk stratification and prognostication in cirrhosis.

This study has several strengths and limitations that warrant attention. The use of the ACE score to enhance the prediction of liver-related outcomes alongside MELD-Na in a large, single-center, longitudinal cohort provided valuable insights. The replication of these findings in a prospective cohort of over 400 waitlisted patients across three Mayo Clinic transplant centers in Minnesota, Arizona, and Florida, coupled with an external validation cohort from Spain, highlighted the potential role of the ACE score in refining organ allocation strategies. During external validation, we observed a decline in performance metrics such as hazard ratios and concordance values with the same increment of the ACE score. This decline is consistent with the general trend of diminished performance of AI algorithms when applied outside their initial development cohort. Therefore, these results need further validation in diverse, international populations across multiple centers. Additionally, the use of different digital formats for ECG storage poses interoperability challenges across institutions which applies to all ECG-based AI models. Researchers are actively developing methods to standardize ECG file storage and to facilitate effective conversion between different digital ECG formats, including DICOM, SCP, and HL7.[28]

The retrospective cohort selection and EHR data extraction were inherently limited by their retrospective nature. Despite unavoidable variations, thorough manual chart reviews were performed to capture the most complete laboratory data and relevant clinical documentation.

Although we used time-updated parameters in our Cox models to capture evolving patient status, we recognize that practical adoption of this approach will require workflow adjustments and possibly automated data retrieval within EHR. In the interim, our results demonstrate the potential value of repeated measurements in refining risk prediction, offering a foundation for prospective studies exploring real-world applicability.

Another notable limitation is the 'black-box' nature of the CNN architecture used in developing the ACE score. While this architecture is a popular choice for computer vision tasks, it complicates the traceability of model outputs to the specific ECG features that clinicians typically recognize. To address this limitation and provide greater transparency, we generated gradient-based saliency maps of representative ECG tracings. These maps revealed that the model predominantly focuses on the QRS complex (and sometimes T wave segments), aligning with the notion that subtle conduction or repolarization changes may be more pronounced in advanced liver disease. Interestingly, in the original study on ACE model development, the performance remained consistent when the ECG length was reduced from 10 s to 2 s, suggesting that the model relies on waveform morphology rather than extended time-dependent patterns such as heart rate.[17] Future research should continue to use these explainability methods to elucidate which specific ECG features—such as intervals, wave amplitudes, or repolarization abnormalities—are most relevant to cirrhosis severity.

In conclusion, the ACE score demonstrates strong predictive potential for liver disease severity and liver-related mortality in patients with cirrhosis. Our findings suggest it could serve as a valuable adjunct to existing prognostic tools, particularly in enhancing risk stratification alongside the MELD score, especially among patients with lower MELD scores. The non-invasive nature of the ACE score and its derivation from widely available ECGs position it as a promising tool for improving disease severity assessment and potentially refining organ allocation models. However, to fully realize its clinical potential, further research is essential. Future studies should focus on developing and validating comprehensive predictive models that integrate the ACE score with other established predictors of liver-related outcomes, conducting explainability studies to understand ECG characteristics influencing the ACE score, and evaluating its performance across diverse populations. Continued investigation is crucial to establish the ACE score's full effectiveness and reliability in clinical practice, ultimately aiming to improve risk stratification, guide treatment decisions, and enhance outcomes for patients with cirrhosis.

Affiliations

[1]*Department of Medicine, Division of Gastroenterology and Hepatology, Mayo Clinic, Rochester, MA, USA;* [2]*Department of Surgery, Division of Hepatobiliary and Pancreas Surgery, Mayo Clinic, Rochester, MA, USA;* [3]*Liver Unit, Hospital Clínic, Barcelona, Catalonia, Spain;* [4]*Institut d'Investigacions Biomèdiques August Pi-Sunyer (IDIBAPS), Barcelona, Catalonia, Spain;* [5]*Centro de Investigaciones Biomédicas en Red de Enfermedades Hepáticas y Digestivas (CIBEREHD), Madrid, Spain;* [6]*Faculty of Medicine and Health Sciences, Barcelona, Catalonia, Spain;* [7]*Department of Medicine, Division of Gastroenterology and Hepatology, Mayo Clinic, Phoenix, AZ, USA;* [8]*Department of Transplantation, Mayo Clinic, Jacksonville, FL, USA;* [9]*Division of Biomedical Statistics and Informatics, Mayo Clinic, Rochester, MN, USA;* [10]*Department of Cardiovascular Medicine, Mayo Clinic, Rochester, MN, USA;* [11]*Department of Surgery, Division of Transplantation Surgery, Mayo Clinic, Rochester, MN, USA*

Abbreviations

ACE, AI-Cirrhosis-ECG; AI, artificial intelligence; CNN, convolutional neural network; EHR, electronic health record; HR, hazard ratio; LRT, likelihood ratio test; LT, liver transplantation; MASLD, metabolic dysfunction-associated steatotic liver disease; MELD, model for end-stage liver disease; MELD-Na score, model for end-stage liver disease-sodium score; OR, odds ratio; ROC, receiver-operating characteristic.

Financial support

DAS and VHS's research is funded by National Institute of Health U01AA026886–03. PG is supported by ISCIII Subdirección General de Evaluación and European Regional Development Fund for the Plan Nacional I+D+I (grant number PI20/00579 del Fondo de Investigaciones Sanitarias) and the Asociación Española para el Estudio del Hígado (AEEH, Beca José Hernandez-Guio).

Conflicts of interest

No other potential conflicts of interest relevant to this article exist.

Please refer to the accompanying ICMJE disclosure forms for further details.

Authors' contributions

Devised the project and the main conceptual ideas for the study: DAS. Performed data extraction: JCA, PR, PS, AJ. Developed, validated, and tested the deep neural network: KL, ZIA. Performed statistical analysis and generated tables and figures: JCA, PR, AFM. Provided patient and ECG data: AJ, MJM, JC, BA, APK. Drafted the manuscript: JCA, PR. Interpreted the results: all authors. Revised the manuscript critically for important intellectual content: PS, AMA, PSK, JKH, PAF, VHS, PAN, PG, DAS. Approved the final version to be published: all authors.

Data availability statement

The data that support the findings of this study are stored within the internal electronic health records systems at Mayo Clinic and Hospital Clínic de Barcelona. These data are not publicly available because of patient privacy considerations. Access to the data is restricted to protect patient confidentiality and comply with institutional policies governing patient health information.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhepr.2025.101356.

References

Author names in bold designate shared co-first authorship

[1] Moon AM, Singal AG, Tapper EB. Contemporary epidemiology of chronic liver disease and cirrhosis. Clin Gastroenterol Hepatol 2020;18:2650–2666.

[2] 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;5:245–266.

[3] Garcia-Tsao G, Friedman S, Iredale J, et al. Now there are many (stages) where before there was one: in search of a pathophysiological classification of cirrhosis. Hepatology 2010;51:1445–1449.

[4] Ginès P, Krag A, Abraldes JG, et al. Liver cirrhosis. Lancet 2021;398:1359–1376.

[5] D'Amico G, Garcia-Tsao G, Pagliaro L. Natural history and prognostic indicators of survival in cirrhosis: a systematic review of 118 studies. J Hepatol 2006;44:217–231.

[6] de Franchis R. Updating consensus in portal hypertension: report of the Baveno III Consensus Workshop on definitions, methodology and therapeutic strategies in portal hypertension. J Hepatol 2000;33:846–852.

[7] de Franchis R, Bosch J, Garcia-Tsao G, et al. Baveno VII - renewing consensus in portal hypertension. J Hepatol 2022;76:959–974.
[8] Kamath PS, Wiesner RH, Malinchoc M, et al. A model to predict survival in patients with end-stage liver disease. Hepatology 2001;33:464–470.
[9] Kamath PS, Kim WR. The model for end-stage liver disease (MELD). Hepatology 2007;45:797–805.
[10] Kim WR, Biggins SW, Kremers WK, et al. Hyponatremia and mortality among patients on the liver-transplant waiting list. N Engl J Med 2008;359:1018–1026.
[11] Lai JC, Covinsky KE, Dodge JL, et al. Development of a novel frailty index to predict mortality in patients with end-stage liver disease. Hepatology 2017;66:564–574.
[12] Starlinger P, Ahn JC, Mullan A, et al. The addition of C-reactive protein and von Willebrand factor to model for end-stage liver disease-sodium improves prediction of waitlist mortality. Hepatology 2021;74:1533–1545.
[13] Kim WR, Mannalithara A, Heimbach JK, et al. Meld 3.0: the model for end-stage liver disease updated for the modern era. Gastroenterology 2021;161:1887. 95.e4.
[14] Asrani SK, Jennings LW, Kim WR, et al. MELD-GRAIL-Na: glomerular filtration rate and mortality on liver-transplant waiting list. Hepatology 2020;71:1766–1774.
[15] Bosch J, Groszmann RJ, Shah VH. Evolution in the understanding of the pathophysiological basis of portal hypertension: how changes in paradigm are leading to successful new treatments. J Hepatol 2015;62:S121–S130.
[16] Solà E, Ginès P. Renal and circulatory dysfunction in cirrhosis: current management and future perspectives. J Hepatol 2010;53:1135–1145.
[17] **Ahn JC**, **Attia ZI**, Rattan P, et al. Development of the AI-Cirrhosis-ECG Score: an electrocardiogram-based deep learning model in cirrhosis. Am J Gastroenterol 2022;117:424–432.
[18] Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care 2005;43:1130–1139.
[19] **Ahn JC**, **Connell A**, Simonetto DA, et al. Application of artificial intelligence for the diagnosis and treatment of liver diseases. Hepatology 2021;73:2546–2563.
[20] Dong TS, Kalani A, Aby ES, et al. Machine learning-based development and validation of a scoring system for screening high-risk esophageal varices. Clin Gastroenterol Hepatol 2019;17:1894. 901.e1.
[21] Liu Y, Ning Z, Örmeci N, et al. Deep convolutional neural network-aided detection of portal hypertension in patients with cirrhosis. Clin Gastroenterol Hepatol 2020;18:2998–3007.e5.
[22] Bosch J, Chung C, Carrasco-Zevallos OM, et al. A machine learning approach to liver histological evaluation predicts clinically significant portal hypertension in NASH cirrhosis. Hepatology 2021;74:3146–3160.
[23] Attia ZI, Kapa S, Lopez-Jimenez F, et al. Screening for cardiac contractile dysfunction using an artificial intelligence–enabled electrocardiogram. Nat Med 2019;25:70–74.
[24] Attia ZI, Noseworthy PA, Lopez-Jimenez F, et al. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. The Lancet 2019;394:861–867.
[25] Ko WY, Siontis KC, Attia ZI, et al. Detection of hypertrophic cardiomyopathy using a convolutional neural network-enabled electrocardiogram. J Am Coll Cardiol 2020;75:722–733.
[26] Galloway CD, Valys AV, Shreibati JB, et al. Development and validation of a deep-learning model to screen for hyperkalemia from the electrocardiogram. JAMA Cardiol 2019;4:428–436.
[27] Attia ZI, Friedman PA, Noseworthy PA, et al. Age and sex estimation using artificial intelligence from standard 12-lead ECGs. Circ Arrhythm Electrophysiol 2019;12:e007284.
[28] Cuevas-González D, García-Vázquez JP, Bravo-Zanoguera M, et al. ECG standards and formats for interoperability between mHealth and healthcare information systems: a scoping review. Int J Environ Res Public Health 2022;19:11941.

Keywords: Artificial intelligence; Deep learning; Hepatic decompensation; Liver transplant; Mortality.
Received 15 May 2024; received in revised form 30 January 2025; accepted 7 February 2025; Available online 19 February 2025