



Published in final edited form as:

J Crit Care. 2012 December ; 27(6): 740.e1–740.e7. doi:10.1016/j.jcrc.2012.08.017.

Predicting In-Hospital Mortality among Critically-Ill Patients with End-Stage Liver Disease

Alex A. Balekian, MD, MSHS¹ and Michael K. Gould, MD, MS²

¹Division of Pulmonary and Critical Care, Keck School of Medicine, University of Southern California, Los Angeles, CA

²Department of Research and Evaluation, Kaiser Permanente Southern California, Pasadena, CA

Abstract

Purpose—Critically-ill patients with end-stage liver disease (ESLD) are at high risk for death during intensive care unit (ICU) hospitalization, and currently available prognostic models have limited accuracy in this population. We aimed to identify variables associated with in-hospital mortality among critically-ill ESLD patients, and to develop and validate a simple, parsimonious model for bedside use.

Materials and Methods—We performed a retrospective chart review of 653 ICU admissions for ESLD patients; modeled in-hospital mortality using multivariable logistic regression; and compared the predictive ability of several different models using the area under receiver operating characteristic (AU-ROC) curves.

Results—Multivariable predictors of in-hospital mortality included Model for End-stage Liver Disease (MELD) score, Acute Physiology and Chronic Health Evaluation (APACHE) II score, mechanical ventilation, and gender; there was also an interaction between MELD score and gender ($p < 0.02$). MELD alone had better discrimination (AU-ROC 0.83) than APACHE II alone (AU-ROC 0.76), and adding mechanical ventilation to MELD achieved the single largest increase in model discrimination (AU-ROC 0.85; $p < 0.01$). In a parsimonious, two-predictor model, higher MELD scores (OR 1.14 per 1-point increase; 95% CI 1.11 – 1.16), and mechanical ventilation (OR 6.20; 95% CI 3.05 – 12.58) were associated with increased odds of death. Model discrimination was also excellent in the validation cohort (AU-ROC 0.90).

Conclusions—In critically-ill ESLD patients, a parsimonious model including only MELD and mechanical ventilation is more accurate than APACHE II alone for predicting in-hospital

© 2012 Elsevier Inc. All rights reserved.

Correspondence should be addressed to: Alex A. Balekian, MD, MSHS, USC Keck School of Medicine, Division of Pulmonary and Critical Care Medicine, 2020 Zonal Avenue, IRD 723, Los Angeles, CA, 90033. Alex.Balekian@med.usc.edu. Phone 818-226-5858; Fax 323-226-2798.

Author Disclosure: Neither author has any financial relationship with a commercial entity that has an interest in the subject of this manuscript. Dr. Balekian was supported by NIH grant 5KL2RR031991 from the University of Southern California's Center for Education, Training, and Career Development (CETCD). Dr. Gould was supported by a Comprehensive Cancer Center Core Support Grant from NIH/NCI (P30 CA014089-34S4).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

mortality. This simple bedside model can provide clinicians and patients with valuable prognostic information for medical decision-making.

Introduction

End-stage liver disease (ESLD) patients who are hospitalized in the intensive care unit (ICU) have mortality rates of 40–80% (1–5). General prognostic models, such as the Acute Physiology and Chronic Health Evaluation (APACHE) II score, have been shown to have limited accuracy when applied to ESLD patients (1,6,7). Liver disease-specific prognostic models, such as the Model for End-stage Liver Disease (MELD) score, were developed to estimate long-term prognosis among patients with chronic liver disease (8,9), with uncertain relevance in the ICU setting.

We performed a retrospective cohort study with two aims: (1) to identify variables associated with in-hospital mortality among critically-ill ESLD patients, and (2) to construct a simple, parsimonious model that could be used at the bedside by clinicians to predict in-hospital mortality in these patients.

Materials and Methods

After obtaining approval from the institutional review board of the University of Southern California (HS-09-00537), we retrospectively reviewed all ICU admissions at Los Angeles County + University of Southern California (LAC+USC) Medical Center, an inner-city, university-affiliated, public teaching hospital with 120 ICU beds (40 medical, 20 neurological, 10 cardiac, 30 surgical, 10 cardiothoracic, and 10 burn) serving a primarily indigent, Hispanic population of Los Angeles County. Using diagnostic codes related to chronic liver disease (ICD-9 codes 571.0 – 571.9), we identified adults (> 18 years) hospitalized between January 2000 and October 2009. We excluded patients whose ICU length of stay (LOS) was less than 1 day. For those patients who were hospitalized multiple times during the study period, we included only the first hospitalization; furthermore, for multiple ICU admissions during a single hospitalization, we included only the first ICU admission. Patients with hepatocellular carcinoma were included in the study, but we excluded those patients with other known malignancies. No patients underwent liver transplantation during hospitalization.

Data Collection and Variable Definitions

We selected potential predictor variables based on literature review and clinical judgment. Variables included demographic characteristics, alcohol use recorded in the physician's admission note, APACHE II score, ICU LOS, hospital LOS, the individual components of the MELD score (creatinine, total bilirubin, international normalized ratio [INR]), mechanical ventilation, and reason for hospitalization (sepsis coded as 991.90 – 991.92; gastrointestinal bleed coded as 578.0 – 578.9).

Data were abstracted from the hospital's computerized medical records system. APACHE II score was entered by ICU nurses within 24 hours of admission. We calculated the MELD score by using the Mayo Clinic web-based calculator (<http://www.mayoclinic.org/meld/>

mayomodel5.html). For those patients admitted directly to the ICU from the emergency department, we used the laboratory values on the day of admission to calculate the MELD score; for those patients transferred to the ICU from the hospital ward, we used the most abnormal lab values recorded in the week prior to transfer. We chose a one-week period based on the United Network for Organ Sharing's (UNOS) practices, detailed in Section 3.6.4 of their Policy on Allocation of Livers (http://optn.transplant.hrsa.gov/PoliciesandBylaws2/policies/pdfs/policy_8.pdf) This practice would also allow us to control for artificial improvements in MELD score achieved by medical therapies (e.g. dialysis, blood transfusions).

The primary outcome was in-hospital death. We assessed this outcome by looking at discharge location (e.g. morgue). Those patients who were listed as being discharged to hospice were also considered to have met the primary outcome. All other discharge or transfer locations— home, nursing home, rehabilitation facility, another hospital – were counted as a non-event, as was leaving against medical advice.

Model Development

We used logistic regression to predict the primary outcome from a limited number of candidate variables, including age, sex, ethnicity, alcohol use, mechanical ventilation, MELD score, APACHE II score, and admission for sepsis or GI bleed. Each candidate variable was examined first in isolation. Those variables that remained significant at the $p=0.10$ level after adjustment were retained in a multivariable logistic regression model that included first-order interactions as well. These predictors were then placed alone or in combination with one another to form one full and several reduced models. We assessed the discrimination of the full and reduced models by deriving the area under the receiver operating characteristic (AU-ROC) curves. We compared the AU-ROC for the full and reduced models by using the non-parametric method of DeLong (10). The calibration of the full and reduced models was assessed by using the Hosmer-Lemeshow goodness-of-fit statistic.

Missing Data

In our original cohort, 178 patients (18.9%) were missing APACHE II score only, 85 patients (9.0%) were missing MELD score only, and 24 patients (2.6%) were missing both APACHE II and MELD scores. Characteristics of patients with missing data are described in Appendix 1. Missing values were imputed by performing linear regression of APACHE II and MELD on other available variables to achieve complete data for 97.7% of our cohort. The results of our multivariable logistic regression analyses (reported in Appendix 1) were not meaningfully different when repeated with and without imputing values for missing data.

Model Validation

Internal validation was performed with bootstrap resampling (653 observations, 200 iterations) using our model-development cohort to obtain bias-corrected estimates of odds ratios in the appropriate candidate model. In addition, we performed a split-sample internal validation of reduced models that did not contain APACHE II score as a predictor by using a

cohort of patients who were missing APACHE II data only and recalculating the AU-ROC. All analyses were performed with SAS Statistical Software version 9 (SAS Institute, Cary, NC).

Results

We identified 940 unique patients with 1071 ICU admissions of at least one day in length (Figure 1). 653 patients (69.5%) had complete data. Patients were predominantly male and Hispanic, with median age of 50 (Table 1). Alcohol use was reported in 60% of patients. In addition, 80% of patients received mechanical ventilation, while 28% received a diagnosis code for sepsis during their hospitalization.

Within the development cohort, survivors were more likely than non-survivors to be younger, male, white or Hispanic, directly admitted to the ICU, not mechanically ventilated, and with lower MELD score, APACHE II score, and ICU LOS (Table 1). Survivors were also less likely to experience sepsis and renal failure requiring dialysis, but more likely to experience gastrointestinal (GI) bleeding.

Predictors of In-Hospital Mortality

After multivariable adjustment (Table 2), patients who required mechanical ventilation were almost four times more likely to die (OR 3.92; 95% confidence interval [CI] 1.82 – 8.47). Likewise, MELD score (OR 1.13 per 1-point increase; 95% CI 1.10 – 1.16), APACHE II score (OR 1.03 per 1-point increase; 95% CI 1.01 – 1.06), age (OR 1.02; 95% CI 1.00 – 1.04), and female sex (OR 1.54; 95% CI 0.94 – 2.54) were associated with a higher risk of death, while GI bleed was associated with a lower risk of death (OR 0.34; 95% CI 0.13 – 0.91). The results of our analysis did not differ markedly after omitting the six patients (1%) discharged to hospice. Those predictors that met our pre-determined level of significance ($p < 0.10$) in the initial multivariable model were placed alone or in combination to construct various predictive models.

The AU-ROC for the full, six-predictor model and several reduced models are shown in Table 3. Thirteen different models had good discrimination (AU-ROC > 0.80) and good fit (Hosmer-Lemeshow statistic $p > 0.05$). MELD alone had good discrimination for the primary outcome (AU-ROC 0.83; 95% CI 0.80 – 0.86), and it performed better than APACHE II alone (AU-ROC 0.73; $p < 0.01$). The discrimination of the MELD model was improved by the addition of mechanical ventilation (AU-ROC 0.85 vs 0.83; $p < 0.001$) or APACHE II score (AU-ROC 0.85 vs 0.83, $p < 0.01$). Additional predictors did not improve model discrimination. Therefore, the most parsimonious model with the best discrimination was a two-predictor model with MELD and mechanical ventilation. The calibration of this model is shown in Figure 2.

Interaction Effects

We tested first-order interaction effects between predictors in the full model. Only sex and MELD score had a significant interaction, and this interaction term remained statistically significant whether added to the full or reduced models. Figure 3 shows the predicted probability of death (using the parsimonious model) based on MELD score after

stratification by sex and mechanical ventilation. At lower MELD scores, women were more likely to die than men after controlling for mechanical ventilation. These sex differences were not apparent at higher MELD scores. Adding sex and its interaction term with MELD to the parsimonious two-predictor model retained the same discrimination (AU-ROC 0.85) but improved its fit (Hosmer-Lemeshow statistic $p=0.33$ vs $p=0.62$).

Because of the significant interaction, we regressed our parsimonious model separately on men and women (Table 4). MELD score was not as strongly associated with death in women (OR 1.09; 95% CI 1.05 – 1.14) as in men (OR 1.15; 95% CI 1.12 – 1.18), but mechanical ventilation was more strongly associated with death in women (OR 16.9; 95% CI 2.1 – 125.8) than in men (OR 5.38; 95% CI 2.43 – 11.87).

Internal Validation

Using bootstrapping with replacement, we internally validated the parsimonious model to obtain bias-corrected estimates for MELD score (OR 1.13 per 1-point increase; 95% CI 1.11 – 1.16) and mechanical ventilation (OR 5.93; 95% CI 2.71 – 12.96). Using split-sample cross-validation with our validation cohort (Table 4), we found that the parsimonious model had better discrimination not only for the entire group (AU-ROC 0.91), but also for men (AU-ROC 0.92) and women (AU-ROC 0.84) separately.

Discussion

In this study of a large, multi-ethnic cohort of indigent, critically ill patients with ESLD, we found that the MELD score predicts in-hospital mortality during an ICU stay with very good discrimination and calibration. We also found that adding a single clinical parameter – mechanical ventilation – to MELD confers improved discriminative ability of the model. The MELD score was initially developed to predict long-term survival after transjugular intrahepatic portosystemic shunt placement, and prior studies have reported modifications that improved the model's performance for specific indications (e.g. adding serum sodium to MELD to optimize protocols for organ allocation) (11). Although APACHE II captures organ dysfunction from multiple systems, results from our study suggest that simple indicators for respiratory failure (need for mechanical ventilation) as well as hepatic and renal failure (captured by the MELD score) can provide important prognostic information to clinicians and families. For example, a woman with a MELD score of 17 who requires mechanical ventilation has a predicted, in hospital mortality of 27%. Our model can also be used as a disease-specific, quality control tool to compare performance across different hospitals in the care of ESLD patients. Furthermore, our model can be used to compare groups and stratify ESLD patients for clinical trials.

Our paper is also the first to identify sex differences in the outcomes of critically-ill ESLD patients, specifically, that at lower MELD scores, women had higher in-hospital mortality than men. Although we did not control for the underlying cause of ESLD – women often suffer more from primary biliary cirrhosis and non-alcoholic liver disease than men – the inclusion of underlying diagnosis did not appreciably improve the discriminative abilities of the MELD score (8). More interestingly, however, the developers of the MELD score did

not include sex in the model, suggesting that this difference might have always existed but was never examined (12).

The reason for this difference between men and women is unclear. A recent article implicated the MELD score's use of creatinine as a reason for this difference, suggesting that this value in female patients with lower body weights did not accurately reflect the degree of renal dysfunction (13). However, when the authors controlled for glomerular filtration rate as well, their model discrimination did not differ significantly from using MELD alone, and even showed a survival advantage for females.

Strengths of our study include the large sample, availability of reliable information about acute and chronic disease severity, use of sensitivity analysis with and without imputation of missing data, and internal validation with both bootstrapping and split-sample techniques. The AU-ROC was higher in the validation cohort, arguing against overfitting of the model to our development cohort. In addition, the one-week window for laboratory values in the calculation of MELD score for patients transferred from the ward allowed us to remain consistent with UNOS policies, to reflect accurately the severity of illness for each observation's particular episode of care, and to minimize information bias for the development of our model.

Limitations of our study include the retrospective design and the possibility of not having adequately captured all critically ill patients with ESLD. However, our rates for gastrointestinal bleeding and sepsis were similar to or higher than a previous retrospective review of hospitalization diagnoses for over 100,000 ESLD patients, suggesting adequate coding of these diseases (14). In addition, although the ethnic and socioeconomic composition of our population might limit the applicability of our findings, our group's mortality rate was comparable to or lower than that reported in previously-published studies (5,15,16), even in the setting of similar or higher MELD and APACHE II scores, suggesting that the findings may be generalizable to groups with other sociodemographic characteristics. Finally, while roughly 1/3 of our original sample had missing data, those with missing information were similar (with respect to age, sex and mortality) to those without missing data. Accordingly, we accounted for missing data by imputation of values (see Appendix) and showing that the results of our analysis did not change. We believe that the precision (for MELD score) and magnitudes (for both) of the odds ratios argue strongly against the possibility that residual confounding from missing data would affect the results.

We also used APACHE II in lieu of APACHE III or IV. Although the later iterations of APACHE improved the model's predictive ability (17,18), they are proprietary, which limits access to their use by a county hospital with limited resources. That notwithstanding, our goal was not to devise an entirely new prediction model, but rather to construct a model that could be used conveniently and efficiently by clinicians at the bedside. Indeed, over twice as many patients in our database were missing APACHE II as compared to the individual components of the MELD score, likely because of the large number of variables necessary to calculate APACHE II (as well as its later iterations).

Our retrospective review did not distinguish whether ventilated patients were intubated for airway protection during a procedure or for respiratory failure. Further studies using an abnormal PaO₂:FiO₂ ratio or alveolar-arterial oxygen gradient as the reason for intubation might strengthen the association of this risk factor and improve the discrimination of this model.

Our findings confirm and extend the results of a prior study that found that MELD was superior to APACHE II for predicting hospital death in cirrhosis (1). However, our study builds on these prior results by using a sample twice as large, by adding mechanical ventilation as a clinical parameter to MELD, and by validating the predictive model in a split sample. Our results differ from those of a recent, smaller study, which did not find MELD score alone to be superior to a modified SOFA score or other general ICU predictive models (19). In this study, the mean MELD score was 23 (compared with 26 in our study), yet the in-hospital mortality rate was 54% (compared with 37% in our sample); this large difference in mortality despite similar MELD scores might be explained by the authors' calculation of the MELD score at the time of ICU transfer from the hospital wards, where transfusions of plasma could have improved the INR and flattered the patient's MELD score, thereby undermining its predictive ability. Furthermore, the authors did not validate their model.

Conclusion

In summary, a two-predictor model using MELD score and mechanical ventilation accurately predicts in-hospital mortality when applied to critically-ill ESLD patients. Differences in survival by sex appear to exist, and this effect appears to be more pronounced at lower MELD scores. Further studies are necessary to validate this model's superiority over the newer models (APACHE IV, SOFA) in other ethnically-diverse cohorts, in addition to elucidating whether sex differences occur more widely. Once more extensively validated, this two-predictor model could prove to be an easy bedside guide to help physicians counsel their ESLD patients and their families about prognosis in critical illness.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

1. Cholongitas E, Senzolo M, Patch D, Kwong K, Nikolopoulou V, Leandro G, Shaw S, Burroughs AK. Risk factors, Sequential Organ Failure Assessment and Model for End-stage Liver Disease scores for predicting short term mortality in cirrhotic patients admitted to intensive care unit. *Alimentary Pharmacology & Therapeutics*. 2006; 23:883–893. [PubMed: 16573791]
2. Gildea TR, Cook WC, Nelson DR, Aggarwal A, Carey W, Younossi Z, Arroliga A. Predictors of long-term mortality in patients with cirrhosis of the liver admitted to a medical ICU. *Chest*. 2004; 126:1598–603. [PubMed: 15539733]
3. Singh N, Gayowski T, Wagener MM, Marino IR. Outcome of patients with cirrhosis requiring intensive care unit support: prospective assessment of predictors of mortality. *J Gastroenterol*. 1998; 33:73–79. [PubMed: 9497225]
4. Kress JP, Rubin A, Pohlman AS, Hall JB. Outcomes of critically ill patients denied consideration for liver transplantation. *Am J Respir Crit Care Med*. 2000; 162:418–423. [PubMed: 10934063]

5. Aggarwal A, Ong JP, Younossi ZM, Nelson DR, Hoffman-Hogg L, Arroliga A. Predictors of mortality and resource utilization in cirrhotic patients admitted to the medical ICU. *Chest*. 2001; 119:1489–1497. [PubMed: 11348958]
6. Zauner C, Schneeweiss B, Schneider B, Madl C, Klos H, Kranz A, Ratheiser K, Kramer L, Lenz K. Short-term prognosis in critically ill patients with liver cirrhosis: an evaluation of a new scoring system. *Eur J Gastroenterol Hepatol*. 2000 May; 12(5):517–22. [PubMed: 10833094]
7. Zauner CA, Apsner RC, Kranz A, Kramer L, Madl C, Schneider B, Schneeweiss B, Ratheiser K, Stockenhuber F, Lenz K. Outcome prediction for patients with cirrhosis of the liver in a medical ICU: a comparison of the APACHE scores and liver-specific scoring systems. *Intensive Care Med*. 1996; 22:559–63. [PubMed: 8814471]
8. Kamath PS, Wiesner RH, Malinchoc M, Kremers W, Thernau TM, Kosberg CL, D'Amico G, Dickson ER, Kim WR. A model to predict survival in patients with end-stage liver disease. *Hepatology*. 2001; 33(2):464–470. [PubMed: 11172350]
9. Wiesner R, Edwards E, Freeman R, Harper A, Kim R, Kamath P, Kremers W, Lake J, Howard T, Merion R, Wolfe R, Krom R. Model for end-stage liver disease (MELD) and allocation of donor livers. *Gastroenterology*. 2003; 124(1):91–96. [PubMed: 12512033]
10. DeLong ER, DeLong DM. *Biometrics*. 1988; 44(3):837–845. [PubMed: 3203132]
11. Biggins SW, Kim R, Terrault N, Saab S, Balan V, Schiano T, Benson J, Thernau T, Kremers W, Wiesner R, Kamath P, Klintmalm G. Evidence-based incorporation of serum sodium concentration into MELD. *Gastroenterology*. 2006; 130(6):1652–60. [PubMed: 16697729]
12. Malinchoc M, Kamath PS, Gordon FD, Peine C, Rank J, ter Borg T. A model to predict poor survival in patients undergoing transjugular intrahepatic portosystemic shunts. *Hepatology*. 2000; 31(4):864–71. [PubMed: 10733541]
13. Myers RP, Shaheen AM, Aspinall AI, Quinn RR, Burak KW. Sex, renal function, and outcomes on the liver transplant waiting list: Assessment of revised MELD including estimated glomerular filtration rate. *Journal of Hepatology*. 2011; 54(3):462–470. [PubMed: 21109324]
14. Foreman M, Mannino DM, Moss M. Cirrhosis as a risk factor for sepsis and death: Analysis of the national hospital discharge survey. *Chest*. 2003; 124 (3):1016–1020. [PubMed: 12970032]
15. Rabe C, Schmitz V, Paasha M, Musch A, Zickermann H, Dumoulin FL, Sauerbruch T, Caselmann WH. Does intubation really equal death in cirrhotic patients? Factors influencing outcome in patients with liver cirrhosis requiring mechanical ventilation. *Intensive Care Med*. 2004; 30:1564–71. [PubMed: 15292984]
16. Thompson SJ, Moran C, Cowan ML, Musa S, Beale R, Treacher D, Hamilton M, Grounds RM, Rahman TM. Outcome of Critically Ill Patients with Cirrhosis Admitted to Intensive Care: An Important Perspective from the Non-transplant Setting. *Alimentary Pharmacology & Therapeutics*. 2010; 32(2):233–243. [PubMed: 20456304]
17. Knaus WA, Wagner DP, Draper EA, Wagner D, Zimmerman J. The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest*. 1991; 100 (6):1619–36. [PubMed: 1959406]
18. Zimmerman JE, Kramer AA, McNair DS, Malila FM. APACHE IV: Hospital mortality assessment for today's critically ill patients. *Crit Care Med*. 2006; 34:1297–1310. [PubMed: 16540951]
19. Das V, Boelle PY, Galbois A, Guidet B, Maury E, Carbonell N, Moreau R, Offenstadt G. Cirrhotic patients in the medical intensive care unit: Early prognosis and long-term survival. *Crit Care Med*. 2010; 38(11):2108–16. [PubMed: 20802324]

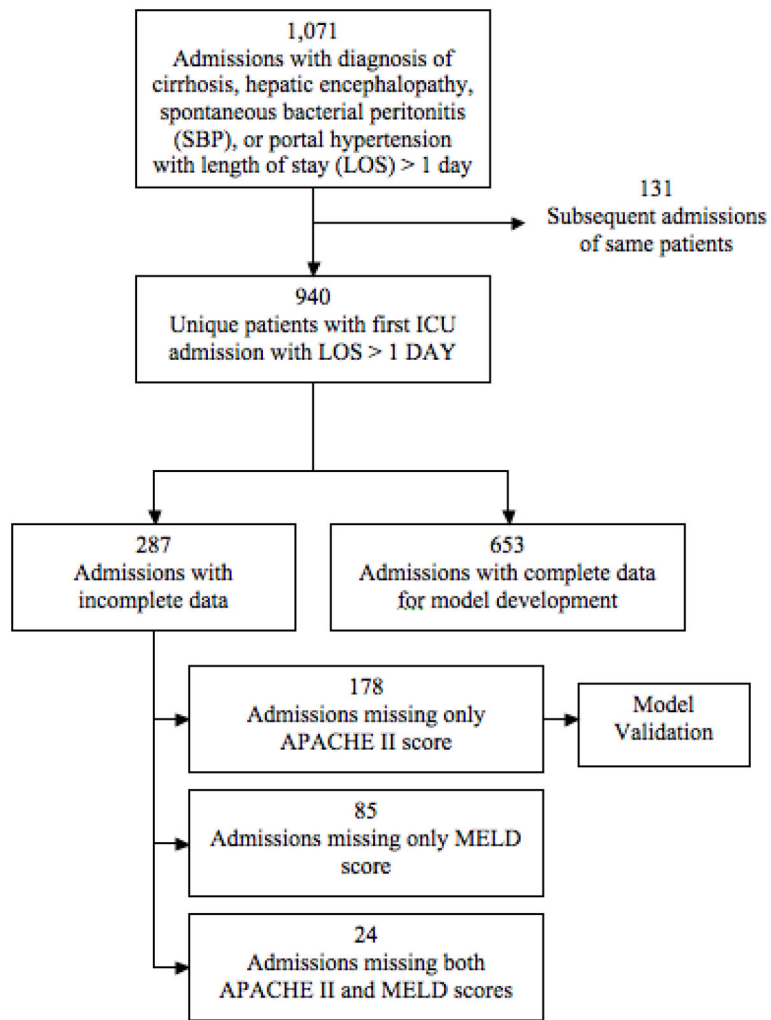


Figure 1. Flow diagram for patient identification and cohorts for development and validation.

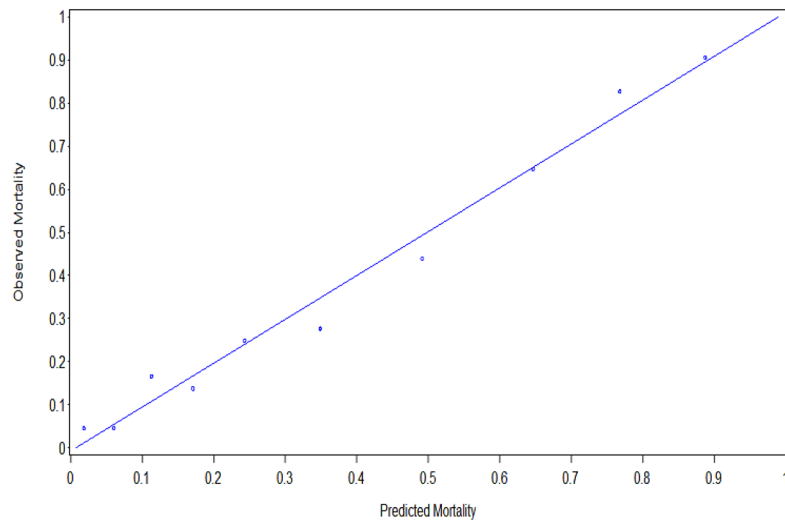


Figure 2. Calibration curve of the parsimonious model. Each point represents one decile of the development cohort.

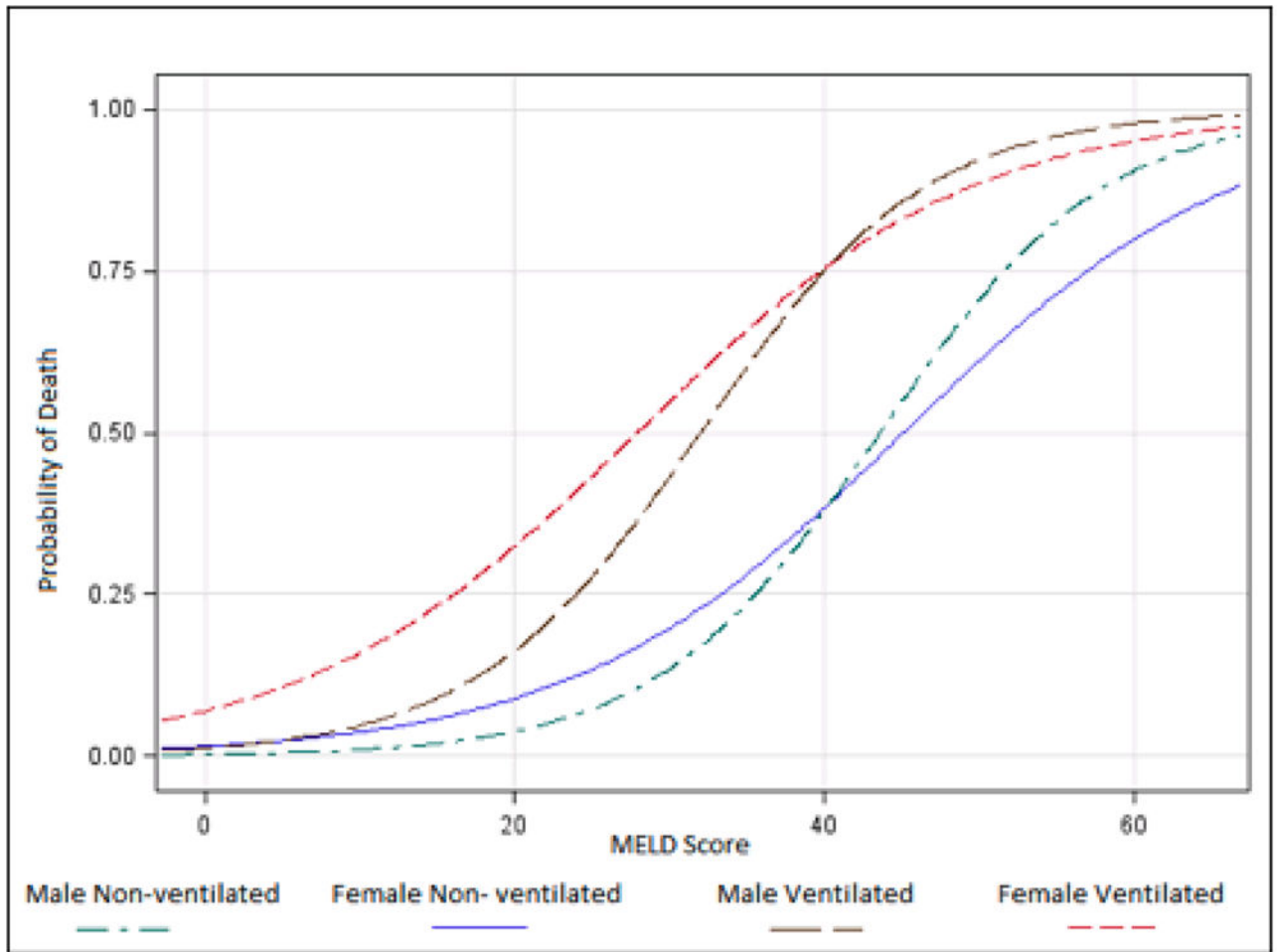


Figure 3. Predicted probability of death based on MELD score when stratifying for sex and mechanical ventilation.

Table 1

Baseline characteristics of survivors and non-survivors in model-development group (n=653). Categorical data presented as number (percentage) and tested with Fisher's exact test; continuous data presented as median (interquartile range) and tested with Wilcoxon rank-sum test.

	All Patients n=653	Survivors n=411 (63%)	Non-survivors n=242 (37%)	p-value
Demographics				
Male	494 (76%)	320 (78%)	174 (72%)	0.09
Age in years	50 (44–57)	50 (44–57)	51 (44–60)	0.04
Race				
White non-Hispanic	62 (9%)	34 (8%)	28 (12%)	0.17
Black	49 (8%)	30 (7%)	19 (8%)	0.88
Hispanic	493 (76%)	320 (78%)	173 (71%)	0.07
Clinical Parameters				
APACHE II	26 (18–33)	22 (16–29)	32 (26–39)	<0.01
MELD	26 (18–37)	21 (15–28)	38 (29–44)	<0.01
MV	523 (80%)	293 (71%)	230 (95%)	<0.01
Dialysis	117 (18%)	38 (9%)	79 (33%)	<0.01
Diagnosis				
Sepsis	180 (28%)	100 (24%)	80 (33%)	0.02
GI Bleed	37 (6%)	29 (7%)	8 (3%)	0.05
Alcohol use	389 (60%)	258 (61%)	138 (57%)	0.32
Direct ICU Admit	456 (70%)	312 (76%)	144 (60%)	<0.01
Outcomes				
ICU LOS in days	6 (3–10)	5 (3–9)	7 (3–13)	<0.01
Home discharge	273 (42%)			
Hospital transfer	60 (9%)			
SNF discharge	51 (8%)			
Left AMA	25 (4%)			
Hospice	6 (1%)			
Expired	236 (36%)			

APACHE II = Acute Physiology and Chronic Health Evaluation II score; MELD = Mayo End-stage Liver Disease score; MV=mechanical ventilation; GI Bleed = gastrointestinal bleed; ICU = intensive care unit; LOS = length of stay; SNF = skilled nursing facility; AMA = against medical advice

Table 2

Results of univariate and multivariate logistic regression for odds ratio (OR) of death using model-development cohort (n=653)

Variable	Unadjusted OR (95% CI)	Adjusted OR (95% CI)	p-value
Age (per year)	1.02 (1.00 – 1.03)	1.02 (1.00 – 1.04)	0.08
Female	1.37 (0.96 – 1.98)	1.54 (0.94 – 2.54)	0.09
Race/ethnicity*			
Black	0.77 (0.36 – 1.65)	1.26 (0.48 – 3.31)	0.82
Hispanic	0.66 (0.39 – 1.12)	0.82 (0.41 – 1.65)	0.11
Clinical Parameters			
MELD**	1.14 (1.12 – 1.16)	1.13 (1.10 – 1.16)	<0.01
APACHE II**	1.11 (1.09 – 1.13)	1.03 (1.01 – 1.06)	0.02
MV	7.72 (4.16 – 14.33)	3.92 (1.82 – 8.47)	<0.01
Diagnosis			
Sepsis	1.54 (1.08 – 2.18)	1.27 (0.79 – 2.04)	0.32
GI Bleed	0.45 (0.20 – 1.00)	0.34 (0.13 – 0.91)	0.03
Alcohol Use	0.85 (0.61 – 1.17)	1.03 (0.65 – 1.63)	0.89

* White non-Hispanic race was used as the baseline comparator

** Odds ratios reported for each 1-point increase

APACHE II = Acute Physiology and Chronic Health Evaluation II score; MELD = Mayo End-stage Liver Disease score; MV=mechanical ventilation; GI Bleed = gastrointestinal bleed

Table 3

Comparison of area under receiver operating characteristic (AU-ROC) curves using model-development cohort

Model	AU-ROC	X ² goodness-of-fit (p-value)
Full Multivariate		
MELD-MV-Sex-APACHE II-Age-GI Bleed	0.86	0.07
Reduced Multivariate		
MELD-MV-Sex-APACHE II-Age	0.86	0.24
MELD-MV-Sex-APACHE II-GI Bleed	0.86	0.16
MELD-MV-Sex-APACHE II	0.86	0.10
MELD- MV-Sex	0.86	0.48
MELD- MV-APACHE II	0.86	0.07
MELD-APACHE II-Sex	0.85	0.51
APACHE II- MV-Sex	0.77	0.56
MELD- MV	0.85	0.33
MELD-APACHE II	0.85	0.38
MELD-Sex	0.84	0.58
MELD-Age	0.84	0.91
MELD-GI Bleed	0.83	0.27
APACHE II- MV	0.77	0.39
APACHE II-Sex	0.76	0.86
APACHE II-GI Bleed	0.76	0.82
APACHE II-Age	0.76	0.83
Bivariate		
MELD	0.83	0.58
APACHE II	0.76	0.66
MV	0.62	--
Age	0.55	0.87
Sex	0.53	--
GI Bleed	0.52	--

APACHE II = Acute Physiology and Chronic Health Evaluation II score; MELD = Mayo End-stage Liver Disease score; MV=mechanical ventilation; GI Bleed = gastrointestinal bleed

Table 4

Split-sample validation results using model-validation cohort (n=178; 133 men and 45 women); Odds ratio of death (95% confidence interval) for each 1-point increase in MELD reported for model-development cohort (n=653; 494 men and 159 women).

Variable	Model-Development Cohort	Model-Validation Cohort
<i>All patients</i>		
MELD	1.14 (1.11 – 1.16)	
Ventilation	6.20 (3.05 – 12.58)	
	AU-ROC 0.85	AU-ROC 0.91
<i>Men</i>		
MELD	1.15 (1.12 – 1.18)	
Ventilation	5.38 (2.43 – 11.87)	
	AU-ROC 0.87	AU-ROC 0.92
<i>Women</i>		
MELD	1.09 (1.05 – 1.14)	
Ventilation	16.90 (2.06 – 125.8)	
	AU-ROC 0.83	AU-ROC 0.84