Clinical Epidemiology

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Development and Validation of a Novel Tool to Predict Model for End-Stage Liver Disease (MELD) Scores in Cirrhosis, Using Administrative Datasets

Tracey G Simon[1,2], Sebastian Schneeweiss[2], Richard Wyss[2], Zhigang Lu[2], Lily G Bessette[2], Cassandra York[2], Kueiyu Joshua Lin[1,2]

[1]Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA; [2]Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

Correspondence: Kueiyu Joshua Lin, Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, 1620 Tremont St. Suite 3030, Boston, MA, 02120, USA, Tel +1 617 278-0930, Fax +1 617 232-8602, Email jklin@bwh.harvard.edu

Background: The Model for End-Stage Liver Disease (MELD) score predicts disease severity and mortality in cirrhosis. To improve cirrhosis phenotyping in administrative databases lacking laboratory data, we aimed to develop and externally validate claims-based MELD prediction models, using claims data linked to electronic health records (EHR).

Methods: We included adults with established cirrhosis in two Medicare-linked EHR networks (training and internal validation; 2007–2017), and a Medicaid-linked EHR network (external validation; 2000–2014). Using least absolute shrinkage and selection operator (LASSO) with 5-fold cross-validation, we selected among 146 investigator-specified variables to develop models for predicting continuous MELD and relevant MELD categories (MELD<10, MELD≥15 and MELD≥20), with observed MELD calculated from laboratory data. Regression coefficients for each model were applied to the validation sets to predict patient-level MELD and assess model performance.

Results: We identified 4501 patients in the Medicare training set (mean age 75.1 years, 18.5% female, mean MELD=13.0), and 2435 patients in the Medicare validation set (mean age: 74.3 years, 31.7% female, mean MELD=12.3). Our final model for predicting continuous MELD included 112 variables, explaining 58% of observed MELD variability; in the Medicare validation set, the area-under-the-receiver operating characteristic curves (AUC) for MELD<10 and MELD≥15 were 0.84 and 0.90, respectively; the AUC for the model predicting MELD≥20 (using 27 variables) was 0.93. Overall, these models correctly classified 77% of patients with MELD<10 (95% CI=0.75–0.78), 85% of patients with MELD≥15 (95% CI=0.84–0.87), and 87% of patients with MELD≥20 (95% CI=0.86–0.88). Results were consistent in the external validation set (n=2240).

Conclusion: Our MELD prediction tools can be used to improve cirrhosis phenotyping in administrative datasets lacking laboratory data.

Keywords: cirrhosis, phenotyping, administrative data, claims

Introduction

Over the past decade, the prevalence of cirrhosis has doubled in the US,[1,2] and rates of cirrhosis-related hospitalizations and mortality are projected to triple by the year 2030.[3] For patients with cirrhosis, disease prognosis varies widely, and depends upon clinical factors including the underlying etiology of liver disease as well as cirrhosis severity. It has also been observed that the safety and treatment efficacy of common medications also may vary according to the severity of liver disease, such that certain common medications – including β-blockers,[4] anti-diabetic agents[5] or oral anticoagulants[6] – may provide less benefit or even increase harm in patients with more advanced, decompensated cirrhosis. The Model for End Stage Liver Disease (MELD) score is an established biomarker of cirrhosis severity that predicts short-term mortality, and it has been clearly demonstrated that clinical outcomes of cirrhotic patients with very low MELD scores differ markedly from those with elevated MELD scores.[7,8] For this reason, the MELD score continues to be widely used for prognostication, including to help guide organ allocation for

Received: 23 August 2022
Accepted: 23 December 2022
Published: 14 March 2023

liver transplantation. Consequently, it is important to carefully account for cirrhosis severity – including MELD score – in clinical studies of patients with cirrhosis.

Large healthcare utilization databases are increasingly used in studies of the natural history of cirrhosis, and for comparative effectiveness research of drug therapies. However, in the field of cirrhosis research, a critical limitation of these datasets is the lack of laboratory information regarding cirrhosis severity, as captured through the MELD score. Specifically, the MELD score may be an important confounder that needs adjustment in non-randomized studies of cirrhotic populations, or it could be a key modifier of other risk factors or observed treatment effects. However, in many large claims datasets, including Medicare and Medicaid, laboratory data are not available. While International Classification of Diseases, Ninth Revision and Tenth Revision (ICD-9 and ICD-10) codes have been proposed to define decompensated cirrhosis, this approach risks underestimating cirrhosis severity, given the frequency with which non-specific codes for cirrhosis are used, in administrative claims datasets.[9]

To date, research focused on improving clinical phenotyping of cirrhosis in administrative claims datasets is very limited. Yet, in other fields, claims-based proxies have demonstrated excellent performance for improving clinical phenotyping of similarly complex conditions, such as heart failure[10,11] and frailty.[12] Thus, we sought to develop and externally validate a claims-based tool for identifying MELD score in patients with cirrhosis, using specific ICD-9 and ICD-10 codes together with important clinical features that differ according to MELD severity.

Methods

Data Source

We used Medicare claims data between 2007 and 2017, from Parts A (inpatient coverage), B (outpatient coverage) and D (prescription drug benefits), and Medicaid data from 2000 to 2014, that included inpatient, outpatient, long-term care, and prescription claims. The Medicare and Medicaid claims data include detailed, patient-level information regarding demographics, enrollment dates, dispensed prescription medications, codes for medical diagnoses and performed procedures. Medicare and Medicaid claims were linked deterministically by date of birth, sex, and health insurance claim numbers or social security number with detailed electronic health record (EHR) data for two large healthcare networks in Boston, Massachusetts (with a linkage success rate of 97.5% in the Medicare population and 98.5% in the Medicaid population). The first network (EHR system 1) consists of 1 tertiary hospital, 2 community hospitals, and 19 primary care centers. The second network (EHR system 2) includes 1 tertiary hospital, 1 community hospital, and 18 primary care centers. The linked EHR data include detailed demographic information, clinical comorbidities, procedures and surgeries, laboratory results, imaging data and histopathology, drug prescribing information as well as reports and notes for all clinical encounters. We used the Medicare data from the first EHR network for model development (training set), and from the second EHR network, we used Medicare data as the internal validation set, and Medicaid data as the external validation set.

The Brigham and Women's Hospital Institutional Review Board approved this study protocol. All accessed data complied with relevant data protection and privacy regulations.

Study Design

The study cohort was required to have valid laboratory results for all three components of the calculated MELD score (including: serum creatinine, bilirubin, and international normalized ratio [INR]) measured within 4 weeks from one another. The index date was defined by the date of the last of the three measured laboratory values. We further required subjects to have at least 6 months of continuous enrollment (specifically: 3 months before and 3 months after the index date, defined as the covariate assessment period [CAP]) in either Medicare Parts A, B and D (the Medicare population) or Medicaid medical and prescription coverage (the Medicaid population). For the Medicare population, all subjects were also required to be 65 years or older, as patients with Medicare who are <65 years constitute a specific, selected population, that is less generalizable than the Medicare population 65 years or older. All subjects were also required to have at least one recorded inpatient or outpatient ICD-9 code [571.2, 571.5, 571.6] or ICD-10 code [K70.3, K74, K74.3, K74.4, K74.5, K74.6, K74.60, K74.69]) for cirrhosis during the CAP. In previous validation studies, analogous ICD-9 and ICD-10 definitions of cirrhosis have yielded positive predictive values (PPV) >85–90%.[13–16]

Outcome Ascertainment

We calculated continuous, observed MELD scores from available laboratory data for each subject, using a validated algorithm.[7,8] We further categorized MELD scores into the following binary variables: (1) MELD<10 vs MELD≥10, as patients with cirrhosis and low MELD<10 have a different natural history and improved prognosis, compared to those with higher MELD; (2) MELD<15 vs MELD≥15, given that a MELD score ≥15 is the clinical threshold at which a patient may be considered eligible for liver transplantation evaluation; and (3) high MELD≥20 vs MELD<20, given the poor prognosis associated with very high MELD scores.[17,18] As patients with very high MELD scores represent the sickest subgroup of patients with cirrhosis, who likely have unique factors contributing to their very high MELD scores, we constructed a separate logistic regression model for this binary variable, as outlined below.

Model Development and Validation

A total of 146 investigator-specified variables were considered candidates for MELD prediction based on their established or putative role in the etiology and natural history of cirrhosis. Briefly, these candidate predictors included demographics (age, sex, race/ethnicity), variables related to etiology of cirrhosis (viral hepatitis B or C infection, nonalcoholic fatty liver disease or steatohepatitis, alcohol-related liver disease or alcoholic hepatitis, other etiologies of chronic liver disease), cirrhosis-specific variables (ICD-9 or 10 codes indicating decompensation events, including ascites/spontaneous bacterial peritonitis, esophageal variceal hemorrhage, hepatic encephalopathy, hepatorenal syndrome or hepatocellular carcinoma, using validated, accepted algorithms for each condition[9]), as well as medication use variables, comorbid conditions and healthcare utilization variables (ie, number of outpatient office visits, number of emergency room visits, number and duration of hospitalizations). Using those candidate predictor variables, we constructed a model to predict continuous MELD score by a linear model with Lasso (least absolute shrinkage and selection operator) regression and 5-fold cross-validation, using the data in our training set, consistent with prior work.[19] To predict the unique subgroup of patients with the most extreme phenotype – who were in the highest MELD category (MELD≥20) - we constructed a separate logistic model with Lasso regression with 5-fold cross-validation, using the data in our training set. For all models, we applied the regression coefficients derived from the Medicare training set to the Medicare internal validation set and the Medicaid external validation set, to predict patient-level MELD scores and assess the performance of each model.

Model Performance

Performance was measured by assessing R from the linear models. We evaluated discrimination by area under the receiver operating characteristic curve (AUC) in logistic regression models for each MELD category (ie [1] MELD<10 vs MELD≥10, [2] MELD≥15 vs MELD<15, and [3] MELD≥20 vs MELD<20). To assess MELD categorization, we ascertained positive predictive values (PPVs, defined as the probability that the predicted MELD category indicated the true MELD category), sensitivity (defined as the probability that a person in a given MELD category would be accurately identified by the model), and overall accuracy, together with exact 95% confidence intervals (CIs). In all cases, the EHR-defined, calculated MELD score was used as the reference standard for the claims-based MELD prediction models. To define the prognostic utility of our MELD prediction tool, we used logistic regression models to evaluate the relationships between predicted MELD scores, measured MELD scores, and the odds of 180-day all-cause mortality, in both the Medicare training and validation sets. Finally, because our two study cohorts had very different age distributions, in a sensitivity analysis, we excluded age from the candidate predictor list to test the robustness of the findings without the need to extrapolate information related to age.

Results

We identified a total of 6936 eligible patients with at least 1 diagnosis of cirrhosis and 6 months of continuous Medicare coverage and available laboratory data for calculating the MELD score, for our training set (EHR network 1; n=4501), internal validation set (EHR network 2; n=2435), and external validation set (the Medicaid population; n=2240; Table S1). As outlined in Table 1, most of the patients in the Medicare training and validation sets were white (85.2% and 89.2%, respectively), and male (81.5% and 68.3%, respectively), with an average age of 75.1 years (standard deviation [SD] 7.4 years) and mean (SD)

Table 1 Patient Characteristics in the Medicare Training Set, Medicare Validation Set, and in the External Medicaid Validation Set

Patient Characteristics	Medicare		Medicaid
	Training Set (N= 4501), N (%)[e]	Internal Validation Set (N=2435), N (%)[e]	External Validation Set (N=2240), N (%)[e]
Age, mean (SD)	75.1 (7.4)	74.3 (7)	48.4 (11.2)
Female	2199 (48.9)	1283 (52.7)	1000 (44.6)
Medicare age categories:	–	–	–
Age: ≥ 90	155 (3.4)	50 (2.1)	–
Age: 85–90	372 (8.3)	187 (7.7)	–
Age: 80–84	614 (13.6)	279 (11.5)	–
Age: 75–79	889 (19.8)	441 (18.1)	–
Age: 70–74	1038 (23.1)	646 (26.5)	–
Age: 65–69	1433 (31.8)	832 (34.2)	–
Medicaid age categories:	–	–	–
Age: ≥ 65	–	–	58 (2.6)
Age: 50–64	–	–	1024 (45.7)
Age: 40–49	–	–	714 (31.9)
Age: 30–39	–	–	266 (11.9)
Age: 18–29	–	–	178 (7.9)
Race / ethnicity	–	–	–
White	4013 (89.2)	2074 (85.2)	1546 (69.0)
Black	151 (3.4)	158 (6.5)	253 (11.3)
Hispanic	69 (1.5)	60 (2.5)	155 (6.9)
Other	213 (4.7)	104 (4.3)	49 (2.2)
Missing	55 (1.2)	39 (1.6)	237 (10.6)
Liver-related comorbidities			
Liver failure	150 (3.3)	75 (3.1)	113 (5)
Alcohol liver disease	500 (11.1)	207 (8.5)	554 (24.7)
Alcohol use disorder	548 (12.2)	230 (9.4)	585 (26.1)
Ascites	3173 (70.5)	1799 (73.9)	1342 (59.9)
Autoimmune liver disease	272 (6)	104 (4.3)	90 (4)
Chronic hepatitis unspecified	91 (2)	32 (1.3)	54 (2.4)
Chronic unspecified liver function	274 (6.1)	86 (3.5)	181 (8.1)
Chronic viral hepatitis	522 (11.6)	201 (8.3)	833 (37.2)
Decompensated cirrhosis	864 (19.2)	413 (17)	636 (28.4)
Esophageal varices (any)	286 (6.4)	94 (3.9)	221 (9.9)

(*Continued*)

Table 1 (Continued).

Patient Characteristics	Medicare		Medicaid
	Training Set (N= 4501), N (%)[e]	Internal Validation Set (N=2435), N (%)[e]	External Validation Set (N=2240), N (%)[e]
Esophageal varices with bleeding	200 (4.4)	79 (3.2)	99 (4.4)
Gastric varices without bleeding	32 (0.7)	9 (0.4)	0 (.)
Hepatocellular carcinoma	344 (7.6)	99 (4.1)	115 (5.1)
Hepatic encephalopathy	386 (8.6)	159 (6.5)	442 (19.7)
Hepatorenal syndrome	68 (1.5)	26 (1.1)	34 (1.5)
Liver transplantation	87 (1.9)	12 (0.5)	37 (1.7)
Nonalcoholic fatty liver disease	473 (10.5)	197 (8.1)	205 (9.2)
Nonalcoholic steatohepatitis	74 (1.6)	20 (0.8)	0 (.)
Paracentesis	987 (21.9)	648 (26.6)	59 (2.6)
Portal hypertension	517 (11.5)	236 (9.7)	241 (10.8)
Primary liver cancer	309 (6.9)	98 (4)	81 (3.6)
Secondary / unspecified cirrhosis	125 (2.8)	58 (2.4)	51 (2.3)
Urgent endoscopy for variceal bleeding	85 (1.9)	16 (0.7)	66 (2.9)
Other comorbidities	–	–	–
Frailty score[a] category: ≥0.34	362 (8)	208 (8.5)	6 (0.3)
Frailty score category: 0.25–0.34	1334 (29.6)	760 (31.2)	257 (11.5)
Frailty score category: 0.15–0.24	2121 (47.1)	1150 (47.2)	1104 (49.3)
Frailty score category: <0.15	684 (15.2)	317 (13)	873 (39)
Combined comorbidity score[b] category: ≥10	1342 (29.8)	908 (37.3)	117 (5.2)
Combined comorbidity score category: 8–9	819 (18.2)	406 (16.7)	225 (10)
Combined comorbidity score category: 4–7	1433 (31.8)	708 (29.1)	775 (34.6)
Combined comorbidity score category: <4	907 (20.2)	413 (17)	1123 (50.1)
MELD[c], mean (SD)	13 (6.5)	12.3 (6.1)	12.3 (6.0)
MELD<10	2032 (45.1)	1220 (50.1)	1066 (47.6)
MELD ≥15	1366 (30.3)	646 (26.5)	572 (25.5)
MELD ≥20	732 (16.3)	364 (14.9)	262 (11.7)
Acute kidney injury	1643 (36.5)	953 (39.1)	436 (19.5)
Atrial fibrillation	1732 (38.5)	911 (37.4)	160 (7.1)

(*Continued*)

Table 1 (Continued).

Patient Characteristics	Medicare		Medicaid
	Training Set (N= 4501), N (%)[e]	Internal Validation Set (N=2435), N (%)[e]	External Validation Set (N=2240), N (%)[e]
Chads Vasc category[d]: ≥6	1678 (37.3)	876 (36)	74 (3.3)
Chads Vasc category: 4–5	1676 (37.2)	918 (37.7)	303 (13.5)
Chads Vasc category: <4	1147 (25.5)	641 (26.3)	1863 (83.2)
Anemia	2642 (58.7)	1475 (60.6)	638 (28.5)
Any gastrointestinal bleed	1316 (29.2)	691 (28.4)	490 (21.9)
Cancer	2388 (53.1)	1456 (59.8)	595 (26.6)
Chronic kidney disease	1661 (36.9)	904 (37.1)	298 (13.3)
Coagulation defects	614 (13.6)	326 (13.4)	230 (10.3)
Chronic obstructive pulmonary disease	1304 (29)	652 (26.8)	304 (13.6)
Coronary revascularization	49 (1.1)	32 (1.3)	8 (0.4)
Dementia	516 (11.5)	274 (11.3)	50 (2.2)
Depression	1368 (30.4)	792 (32.5)	531 (23.7)
Dialysis	316 (7)	167 (6.9)	112 (5)
Drug abuse	248 (5.5)	109 (4.5)	495 (22.1)
Deep vein thrombosis	649 (14.4)	398 (16.3)	165 (7.4)
End stage renal disease	304 (6.8)	160 (6.6)	90 (4)
Falls	834 (18.5)	420 (17.2)	115 (5.1)
Flu vaccine	1076 (23.9)	583 (23.9)	79 (3.5)
Foot ulcer	310 (6.9)	161 (6.6)	73 (3.3)
Gangrene	74 (1.6)	32 (1.3)	30 (1.3)
Gastritis or esophagitis	665 (14.8)	349 (14.3)	269 (12)
Gastroesophageal reflux disease	1569 (34.9)	933 (38.3)	242 (10.8)
Heart Failure	1931 (42.9)	1034 (42.5)	429 (19.2)
Hemochromatosis	77 (1.7)	47 (1.9)	27 (1.2)
Hyperkalemia	715 (15.9)	355 (14.6)	126 (5.6)
Hyperlipidemia	2725 (60.5)	1504 (61.8)	230 (10.3)
Hypertension	3673 (81.6)	2001 (82.2)	719 (32.1)
Hypotension	1592 (35.4)	861 (35.4)	340 (15.2)
Intracranial bleed	152 (3.4)	96 (3.9)	90 (4)
Ischemic heart	1857 (41.3)	1099 (45.1)	302 (13.5)
Ischemic stroke	808 (18)	333 (13.7)	160 (7.1)

(*Continued*)

Table 1 (Continued).

Patient Characteristics	Medicare		Medicaid
	Training Set (N= 4501), N (%)[e]	Internal Validation Set (N=2435), N (%)[e]	External Validation Set (N=2240), N (%)[e]
Late effects of cerebrovascular disease	348 (7.7)	191 (7.8)	69 (3.1)
Lower extremity amputation	75 (1.7)	46 (1.9)	28 (1.3)
Lower gastrointestinal bleed	1219 (27.1)	648 (26.6)	474 (21.2)
Major bleed	304 (6.8)	143 (5.9)	80 (3.6)
Obesity	718 (16)	417 (17.1)	151 (6.7)
Pulmonary embolism	287 (6.4)	200 (8.2)	96 (4.3)
Peptic ulcer disease	2233 (49.6)	1246 (51.2)	537 (24)
Peripheral vascular disease	992 (22)	483 (19.8)	81 (3.6)
Surgical aortic valve replacement	42 (0.9)	23 (0.9)	7 (0.3)
Sleep apnea	438 (9.7)	247 (10.1)	74 (3.3)
Type 1 diabetes mellitus	413 (9.2)	200 (8.2)	144 (6.4)
Type 2 diabetes mellitus	1817 (40.4)	1004 (41.2)	506 (22.6)
Upper endoscopy	1346 (29.9)	687 (28.2)	632 (28.2)
Medication use			
Angiotensin-converting enzyme (ACE) inhibitors	1298 (28.8)	681 (28)	404 (18)
Angiotensin II receptor antagonists (ARBs)	200 (4.4)	112 (4.6)	23 (1)
Antiarrhythmics	201 (4.5)	127 (5.2)	24 (1.1)
Antibiotics	2454 (54.5)	1385 (56.9)	1217 (54.3)
Anti-obesity medications	12 (0.3)	2 (0.1)	0 (.)
Antiplatelets	398 (8.8)	215 (8.8)	142 (6.3)
Apixaban	0 (.)	0 (.)	0 (.)
Betablockers	2550 (56.7)	1355 (55.6)	815 (36.4)
Calcium channel blockers	445 (9.9)	249 (10.2)	112 (5)
COX-2 inhibitors	51 (1.1)	29 (1.2)	39 (1.7)
Dabigatran	25 (0.6)	9 (0.4)	0 (.)
Histamine H2-receptor antagonists	325 (7.2)	222 (9.1)	284 (12.7)
Insulin	348 (7.7)	180 (7.4)	302 (13.5)
Lactulose	397 (8.8)	182 (7.5)	429 (19.2)
Loop diuretics	1869 (41.5)	957 (39.3)	660 (29.5)

(*Continued*)

Table 1 (Continued).

Patient Characteristics	Medicare		Medicaid
	Training Set (N= 4501), N (%)[e]	Internal Validation Set (N=2435), N (%)[e]	External Validation Set (N=2240), N (%)[e]
Non-insulin antidiabetic medications	797 (17.7)	390 (16)	227 (10.1)
Nonselective beta blockers	531 (11.8)	231 (9.5)	296 (13.2)
Non-steroidal anti-inflammatory drugs (NSAIDs)	423 (9.4)	293 (12)	490 (21.9)
Proton pump inhibitors	2062 (45.8)	1136 (46.7)	1089 (48.6)
Rifaximin	125 (2.8)	51 (2.1)	58 (2.6)
Rivaroxaban	62 (1.4)	47 (1.9)	2 (0.1)
Spironolactone	686 (15.2)	299 (12.3)	495 (22.1)
Statin	2007 (44.6)	1090 (44.8)	236 (10.5)
Warfarin	1049 (23.3)	485 (19.9)	189 (8.4)

Notes: [a]Frailty was measured using a claims-based frailty index (CFI) validated in the Medicare population.[12,26–28] [b]The burden of comorbidity was quantified using a combined comorbidity score.[29] [c]For MELD calculation, see Methods. [d]CHADS-Vasc score is a risk score for predicting stroke in patients with atrial fibrillation, and includes age, sex, history of congestive heart failure, hypertension, prior stroke or transient ischemic attack, history of vascular disease and history of diabetes.[30] [e]For any variable with no individuals in a given cohort, the Number (%) for that variable is represented by 0 (.).
Abbreviations: N, number; SD, standard deviation; MELD, model for end-stage liver disease.

MELD score of 13.0 (6.5) in the training set, and average age of 74.3 years (SD 7.0 years) with mean (SD) MELD 12.3 (6.1) in the initial validation set (Table 1). In the Medicaid external validation set, we observed a younger mean age, higher proportion of non-white races and lower comorbidity burden, yet similar MELD distribution (Table 1).

Performance of the MELD Prediction Models

As outlined in Table 2, the mean measured MELD score increased appropriately with increasing values of the claims-based predicted MELD score, from 8.0 (SD 1.7; <10th percentile) to 24.1 (SD 5.4; >90th percentile), and these findings were consistent in both the Medicare internal validation cohort and in the Medicaid external validation cohort. In all 3 cohorts (ie, the Medicare training set, Medicare validation set, and the Medicaid external validation set), the predicted MELD score closely approximated the observed MELD categories, for both low MELD<10 and for MELD≥15 (Tables S2 and S3). As outlined above, we also constructed a separate model for predicting the extreme, most severe phenotype (ie, MELD≥20) which is likely to have unique associations with predictors, and this model also closely approximated the observed MELD≥20 category, in all 3 cohorts (Table S4).

Table 3 and Figure 1 outline the performance of the MELD prediction models in the Medicare training and internal validation sets, and in the external Medicaid validation set. The final model for predicting continuous MELD included 112 readily available claims-based variables, and demonstrated excellent discrimination in the training set, with R^2=0.58 for the linear model of continuous MELD, and AUC 0.86 for predicting MELD<10, and 0.91 for MELD≥15; the final model for predicting MELD≥20 included 27 variables and demonstrated an AUC of 0.93 (Table 3).

Within the Medicare validation set, the logistic models for each MELD category showed appropriate overall accuracy (between 0.75 and 0.96), as well as good sensitivity and PPV for identifying low MELD <10 (sensitivity=0.79, 95% CI=0.77–0.81; PPV=0.76, 95% CI=0.73–0.78), and MELD ≥15 (sensitivity=0.60 [95% CI=0.56–0.64]; PPV=0.80, 95% CI=0.76–0.83). The PPV for identifying very high MELD≥20 was fair (PPV=0.80, 95% CI=0.74–0.85); however, the sensitivity was diminished, due to the small number of subjects in this very-sick subgroup (sensitivity=0.49, 95% CI=0.44–0.54).

Table 2 Mean Observed MELD Score by Deciles of Predicted MELD Score

Deciles by Predicted MELD	Medicare		Medicaid
	Mean Observed MELD (SD) in the Training Set	Mean Observed MELD (SD) in the Internal Validation Set	Mean Observed MELD (SD) in the External Validation Set
0	7.95 (1.68)	8.05 (1.74)	8.16 (2.04)
1	8.28 (2.27)	8.37 (2.32)	9.00 (2.81)
2	8.72 (2.45)	8.58 (2.46)	9.45 (2.88)
3	9.65 (3.24)	9.02 (2.83)	9.86 (3.41)
4	10.64 (3.39)	10.33 (3.41)	10.79 (3.79)
5	12.01 (4.29)	11.31 (4.04)	11.65 (4.45)
6	13.71 (5.02)	12.78 (5.01)	12.52 (4.97)
7	15.95 (5.84)	14.28 (4.97)	13.85 (4.80)
8	18.80 (5.90)	18.24 (6.15)	15.60 (6.50)
9	24.14 (5.42)	22.49 (5.70)	22.03 (6.97)

Abbreviations: MELD, model for end-stage liver disease; SD, standard deviation.

Table 3 Performance of the MELD Prediction Score

Model/ Categories	AUC in Training Set	AUC in Internal Validation Set	AUC in External Validation Set	Performance in Medicare Internal Validation Set*			Performance in Medicaid External Validation Set*		
				Overall Accuracy (95% CI)	PPV (95% CI)	Sensitivity (95% CI)	Overall Accuracy (95% CI)	PPV (95% CI)	Sensitivity (95% CI)
MELD <10	0.86	0.84	0.78	0.76 (0.74, 0.77)	0.81 (0.78, 0.83)	0.68 (0.65, 0.70)	0.71 (0.69, 0.73)	0.68 (0.65, 0.70)	0.73 (0.71, 0.76)
MELD ≥15	0.91	0.90	0.85	0.85 (0.84, 0.86)	0.75 (0.71, 0.79)	0.65 (0.61, 0.68)	0.82 (0.80, 0.84)	0.84 (0.79, 0.88)	0.36 (0.32, 0.40)
MELD ≥20	0.93	0.93	0.88	0.91 (0.89, 0.92)	0.80 (0.74, 0.85)	0.49 (0.44, 0.54)	0.91 (0.90, 0.92)	0.78 (0.70, 0.86)	0.32 (0.26, 0.37)

Notes: *The negative predictive value (NPV) and specificity were also evaluated for each of the MELD prediction models. Within the Medicare internal validation set, the NPVs for predicted MELD<10, ≥15 and ≥20 were, 0.72 (95% CI 0.70–0.74), 0.88 (95% CI 0.86–0.89) and 0.92 (95% CI 0.90–0.93), respectively, with corresponding specificity of, 0.84 (95% CI 0.82–0.86), 0.92 (95% CI 0.91–0.93) and 0.98 (95% CI 0.97–0.98), respectively. Within the Medicaid external validation set, the NPVs for predicted MELD<10, ≥15 and ≥20 were, 0.74 (95% CI 0.71–0.76), 0.82 (95% CI 0.80–0.83), and 0.92 (0.90–0.93), respectively, with corresponding specificity of, 0.68 (95% CI 0.66–0.71), 0.98 (95% CI 0.97–0.98) and 0.99 (95% CI 0.98–0.99), respectively.
Abbreviations: AUC, area under the receiver operating characteristic; CI, confidence interval; PPV, positive predictive value; MELD, model for end stage liver disease.

In the Medicaid external validation set, we observed similar discrimination, with AUC between 0.79 and 0.88 for the three MELD categories, and overall accuracy was appropriate for predicting specific MELD categories (between 0.71 and 0.91; Table 3). The PPVs for each MELD category in the external validation set were, 0.68 (95% CI=0.65–0.71) for MELD<10, 0.84 (95% CI=0.80–0.89) for MELD≥15, and 0.78 (95% CI=0.70–0.86) for MELD≥20. While the sensitivity was adequate for predicting low MELD<10 (0.74, 95% CI=0.72–0.77), given the smaller numbers of subjects in the sicker Medicaid subgroups with moderate and high MELD scores, the sensitivity for predicting MELD≥15 and MELD≥20 was diminished (0.36 [95% CI=0.32–0.40], and 0.32 [95% CI=0.26–0.37], respectively). Table 3 also includes the negative predictive values and specificity for each MELD category.

Supplementary Tables S5 and S6 outline all variables included in the final MELD prediction models together with their coefficients, so that researchers in hepatology can more accurately phenotype cirrhosis severity using the predicted

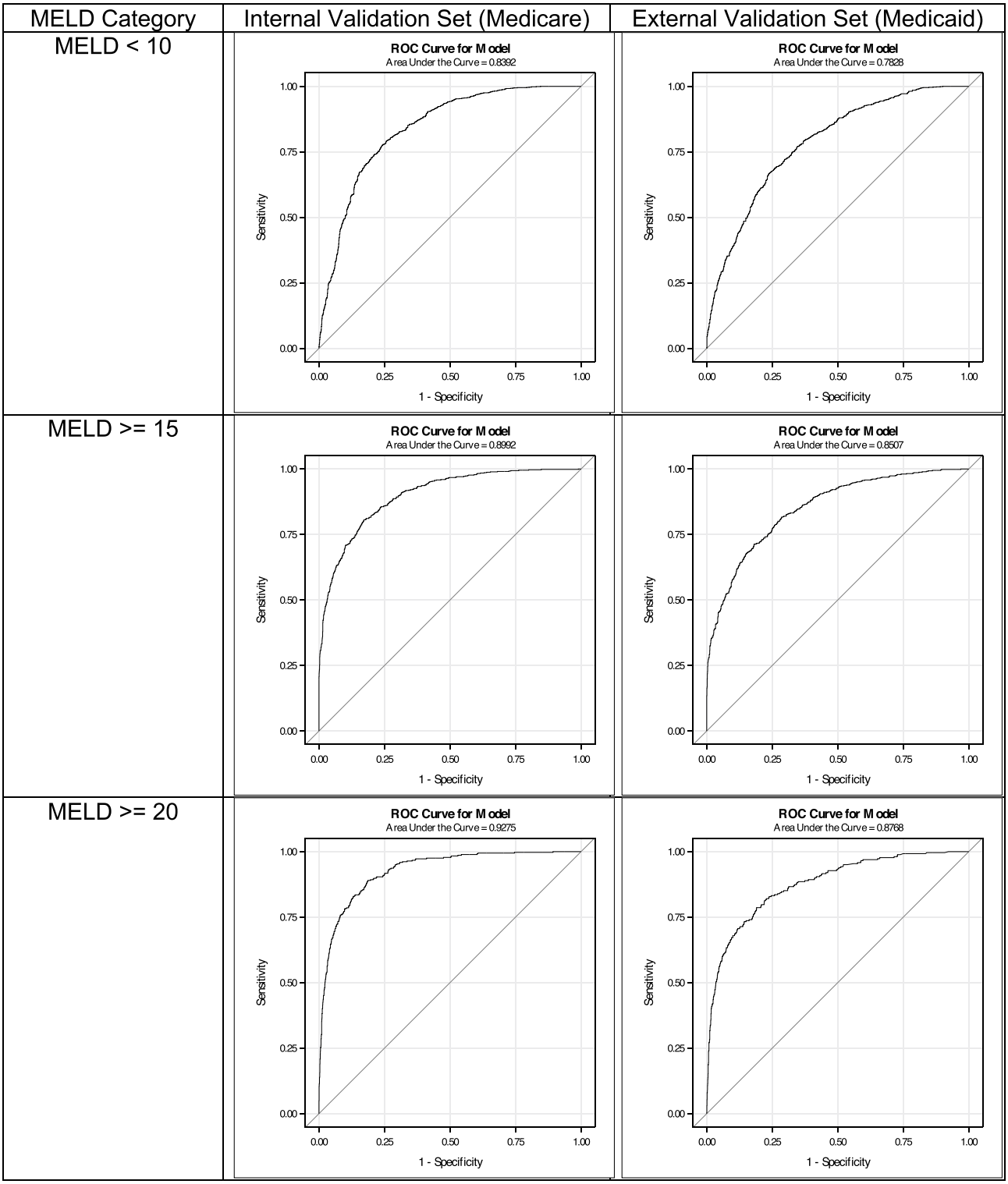

Figure 1 Area under the receiver operating characteristic (AUROC) curves for the performance of the MELD prediction tool in the medicare internal validation set and the Medicaid external validation set.

Abbreviations: ROC, receiver operating characteristic; MELD, model for end-stage liver disease.

MELD score in administrative datasets, when laboratory data to calculate MELD scores are not available. Also, we tested the associations between predicted MELD, measured MELD, and odds of 180-day all-cause mortality, in the Medicare training and validation sets (Supplementary Table S7). In the Medicare training set, compared to low predicted

MELD<10, patients with predicted MELD 10-<20 had 1.4-fold higher odds of 180-day mortality, while patients with predicted MELD ≥20 had 2.7-fold higher odds of 180-day mortality. Importantly, these estimates were similar to those obtained using actual (measured) MELD score categories, and they also were consistent in the validation set (Supplementary Table S7). Finally, in a sensitivity analysis, we excluded age from the model as a candidate predictor and did not find appreciable difference in AUC when predicting all MELD categories (MELD <10, ≥15, or ≥20) in the Medicare training and validation sets as well as the Medicaid validation set (Supplementary Table S8).

Discussion

MELD score is an important indicator of disease severity and major predictor of clinical outcomes, in patients with cirrhosis, thus it is important to develop tools to accurately phenotype cirrhotic populations according to their MELD scores, in large healthcare databases that typically lack sufficient laboratory data. In this study, we developed and validated claims-based models for predicting MELD scores in patients with established cirrhosis. Our findings demonstrate that these models have high overall accuracy for ascertaining MELD scores within both Medicare and Medicaid claims data, and that they are capable of correctly classifying patients with cirrhosis according to clinically relevant MELD categories.

Our model leveraged key differences in the characteristics of cirrhotic patients across the MELD continuum, including age, sex, medication use, clinical comorbidities, and healthcare utilization variables, together with ICD-9 and ICD-10 codes for cirrhosis and its complications, to identify overall MELD score and further to discriminate between low, intermediate, and high MELD score categories. In the final model, the coefficients most strongly and positively linked to MELD prediction included use of warfarin, cardiovascular comorbidities, chronic kidney disease, and variables related to the severity and specific etiologies of liver disease. It is well-established that patients with increasingly severe liver disease are more likely to have chronic kidney disease and cardiovascular disease, as well as an increased likelihood of developing thromboses or arrhythmias, for which anticoagulation might be indicated. Although warfarin does increase the MELD score (by increasing INR), in current clinical practice the MELD score is not "corrected" for warfarin use. Thus, we believe that these findings demonstrate the face validity of our models.

To our knowledge, no prior study in an administrative dataset has attempted to classify patients with cirrhosis into predicted MELD categories using a combination of ICD codes and relevant patient characteristics. While some prior studies have used ICD-9 and ICD-10 codes to identify decompensated cirrhosis, this ICD-only approach is likely to underestimate the actual prevalence of decompensated cirrhosis, and the inclusion of laboratory-based information provided by the MELD score adds substantial additional prognostic information to the assessment of a patient with cirrhosis. Thus, our model provides a practical solution for researchers utilizing large healthcare databases to study cirrhosis, by providing coefficients for the predictor variables – all of which are readily available within claims datasets – to predict continuous MELD score and MELD class. Moreover, the use of independent datasets with known differences in the severity of liver disease and its complications,[20,21] for training and validation, reduced the potential for overfitting.

Cirrhosis represents a rapidly growing cause for hospitalization among Medicare and Medicaid beneficiaries, however the rates of hospitalization and clinical outcomes of these patients vary widely depending on the severity of cirrhosis and other patient-specific factors.[22] As such, our model has potentially important implications for both clinical and health services research, in cirrhosis. First, this model can help reduce potential confounding by severity of cirrhosis, in research studies of cirrhosis outcomes. Second, treatment patterns and their effects vary with cirrhosis severity, thus our model can enable valid assessments of treatment effect heterogeneity according to MELD class, within Medicare and Medicaid participants. Third, hospital performance metrics include rates of readmission and mortality, and administrative claims-based models using public insurance data have been proposed to assess such outcomes, in other diseases.[23,24] Cirrhosis-specific hospital performance metrics have recently been proposed,[25] and by accounting for cirrhosis severity through the MELD score, our model would allow for separate assessments of hospital or clinic performance for cirrhotic patients according to MELD category. Finally, given the high cost of hospitalizations for cirrhosis in Medicare and Medicaid, and the higher prevalence of more advanced liver disease and liver-related hospitalizations in Medicaid, improved risk stratification based on MELD categorization may help to improve resource allocation.

Our model had good PPV in the external validation set for identifying low MELD<10 (76%), MELD≥15 (80%) and MELD≥20 (93%). These estimates, together with the strong overall accuracy (77% to 87%) and excellent discrimination

indicate that this model has appropriate performance for clinical research using Medicare and Medicaid claims data. Of note, the sensitivity for ascertaining MELD≥20 was comparatively modest, owing to the small sample size of this subgroup of very sick patients with the highest MELD scores.

We acknowledge several limitations. First, our samples included Medicare and Medicaid beneficiaries linked to two multicenter Boston metropolitan EHR systems; thus, our results may not be generalizable to other populations from other regions, including those that differ by race/ethnicity or insurance status or the etiology of liver disease, which was unknown in 50% of patients in the derivation set. Second, while EHR data provides rich and detailed clinical information, missing data are inevitable, which could impact generalizability. Third, this study focused on predicted MELD score, rather than the more recently developed MELD-Na score, or even the emerging MELD 3.0; thus, we look forward to future studies that use this framework and approach to develop and validate models for these and other important prognostic variables for cirrhosis outcomes, and to evaluate the correlation between different scoring systems. Finally, our population was limited to participants with recorded laboratory values sufficient to calculate MELD scores, as this served as our reference-standard for MELD ascertainment. Since creatinine, sodium, and INR data are all commonly used, standard laboratory measures, we expect that they are available in most patients with established cirrhosis; thus, our cohort is likely to be representative of typical patients with known cirrhosis in routine clinical care, except for those with undiagnosed cirrhosis, very mild disease, or those with minimal medical encounters. Nevertheless, additional studies are needed in more diverse populations, including in patients with private insurance, uninsured patients, and in large, multi-center EHR networks. Related to this, MELD was assessed using variables collected within 4 weeks of one another, and for some patients with end-stage liver disease, MELD may fluctuate more rapidly, which could lead to inaccuracies in predicting true MELD scores. Thus, it will be important for further research to investigate models constructed using shorter time intervals, particularly for end-stage liver disease.

In conclusion, our novel claims-based tool accurately identifies overall MELD score as well as specific, clinically relevant MELD categories, when the necessary laboratory data for MELD calculation are not available. Thus, our model may be used to better phenotype and risk stratify patients with cirrhosis engaged in routine clinical care, and thereby improve research studies of cirrhosis clinical outcomes and healthcare utilization in administrative databases.

Data Sharing Statement

No additional data are available.

Ethics

The Brigham and Women's Hospital Institutional Review Board approved this study protocol. All accessed data complied with relevant data protection and privacy regulations.

Funding

NIH RO1LM013204 (Lin), NIH K23 DK122104 (Simon), NIH R01HL167021 (Simon). No funding organization had any role in the design and conduct of the study; in the collection, management, and analysis of the data; or in the preparation, review, and approval of the manuscript.

Disclosure

TGS has received research funding from Amgen and has received consulting fees from Aetion, for work unrelated to this manuscript. SS is participating in investigator-initiated grants to the Brigham and Women's Hospital from Boehringer Ingelheim unrelated to the topic of this study. He is a consultant to Aetion Inc., a software manufacturer of which he owns equity. His interests were declared, reviewed, and approved by the Brigham and Women's Hospital and Partners HealthCare System in accordance with their institutional compliance policies. LGB reports personal fees from Amazon Web Services, personal fees from Aetion Inc., outside the submitted work. The authors report no other conflicts of interest in this work.

References

1. Beste LA, Leipertz SL, Green PK, et al. Trends in burden of cirrhosis and hepatocellular carcinoma by underlying liver disease in US veterans, 2001–2013. *Gastroenterology*. 2015;149(6):1471–82e5; quiz e17–8. doi:10.1053/j.gastro.2015.07.056
2. Tapper EB, Parikh ND. Mortality due to cirrhosis and liver cancer in the United States, 1999–2016: observational study. *BMJ*. 2018;362:k2817. doi:10.1136/bmj.k2817
3. Estes C, Razavi H, Loomba R, et al. Modeling the epidemic of nonalcoholic fatty liver disease demonstrates an exponential increase in burden of disease. *Hepatology*. 2018;67(1):123–133. doi:10.1002/hep.29466
4. Rodrigues SG, Mendoza YP, Bosch J. Beta-blockers in cirrhosis: evidence-based indications and limitations. *JHEP Rep*. 2020;2(1):100063. doi:10.1016/j.jhepr.2019.12.001
5. Elkrief L, Rautou PE, Sarin S, et al. Diabetes mellitus in patients with cirrhosis: clinical implications and management. *Liver Int*. 2016;36 (7):936–948. doi:10.1111/liv.13115
6. Qamar A, Vaduganathan M, Greenberger NJ, et al. Oral anticoagulation in patients with liver disease. *J Am Coll Cardiol*. 2018;71(19):2162–2175. doi:10.1016/j.jacc.2018.03.023
7. Kamath PS, Wiesner RH, Malinchoc M, et al. A model to predict survival in patients with end-stage liver disease. *Hepatology*. 2001;33 (2):464–470. doi:10.1053/jhep.2001.22172
8. Wiesner R, Edwards E, Freeman R, et al. Model for end-stage liver disease (MELD) and allocation of donor livers. *Gastroenterology*. 2003;124 (1):91–96. doi:10.1053/gast.2003.50016
9. Hagstrom H, Adams LA, Allen AM, et al. Administrative coding in electronic health care record-based research of NAFLD: an expert panel consensus statement. *Hepatology*. 2021;74(1):474–482. doi:10.1002/hep.31726
10. Desai RJ, Lin KJ, Patorno E, et al. Development and preliminary validation of a medicare claims-based model to predict left ventricular ejection fraction class in patients with heart failure. *Circ Cardiovasc Qual Outcomes*. 2018;11(12):e004700. doi:10.1161/CIRCOUTCOMES.118.004700
11. Mahesri M, Chin K, Kumar A, et al. External validation of a claims-based model to predict left ventricular ejection fraction class in patients with heart failure. *PLoS One*. 2021;16(6):e0252903. doi:10.1371/journal.pone.0252903
12. Kim DH, Schneeweiss S, Glynn RJ, et al. Measuring frailty in medicare data: development and validation of a claims-based frailty index. *J Gerontol a Biol Sci Med Sci*. 2018;73(7):980–987. doi:10.1093/gerona/glx229
13. Bengtsson B, Askling J, Ludvigsson JF, et al. Validity of administrative codes associated with cirrhosis in Sweden. *Scand J Gastroenterol*. 2020:1–6. doi:10.1080/00365521.2020.1820566
14. Goldberg D, Lewis J, Halpern S, et al. Validation of three coding algorithms to identify patients with end-stage liver disease in an administrative database. *Pharmacoepidemiol Drug Saf*. 2012;21(7):765–769. doi:10.1002/pds.3290
15. Mapakshi S, Kramer JR, Richardson P, et al. Positive predictive value of international classification of diseases, 10th revision, codes for cirrhosis and its related complications. *Clin Gastroenterol Hepatol*. 2018;16(10):1677–1678. doi:10.1016/j.cgh.2018.01.042
16. Tapper EB, Korovaichuk S, Baki J, et al. Identifying patients with hepatic encephalopathy using administrative data in the ICD-10 era. *Clin Gastroenterol Hepatol*. 2019. doi:10.1016/j.cgh.2019.12.017
17. Merion RM. When is a patient too well and when is a patient too sick for a liver transplant? *Liver Transpl*. 2004;10(10Suppl 2):S69–S73. doi:10.1002/lt.20265
18. Martin AP, Bartels M, Hauss J, et al. Overview of the MELD score and the UNOS adult liver allocation system. *Transplant Proc*. 2007;39 (10):3169–3174. doi:10.1016/j.transproceed.2007.04.025
19. Lin KJ, Singer DE, Glynn RJ, et al. Prediction score for anticoagulation control quality among older adults. *J Am Heart Assoc*. 2017;6(10). doi:10.1161/JAHA.117.006814
20. Sellers CM, Uhlig J, Ludwig JM, et al. The impact of socioeconomic status on outcomes in hepatocellular carcinoma: inferences from primary insurance. *Cancer Med*. 2019;8(13):5948–5958. doi:10.1002/cam4.2251
21. Robinson A, Hirode G, Wong RJ. Ethnicity and insurance-specific disparities in the model for end-stage liver disease score at time of liver transplant waitlist registration and its impact on mortality. *J Clin Exp Hepatol*. 2021;11(2):188–194. doi:10.1016/j.jceh.2020.07.011
22. Asrani SK, Kouznetsova M, Ogola G, et al. Increasing health care burden of chronic liver disease compared with other chronic diseases, 2004–2013. *Gastroenterology*. 2018;155(3):719–29e4. doi:10.1053/j.gastro.2018.05.032
23. Keenan PS, Normand SL, Lin Z, et al. An administrative claims measure suitable for profiling hospital performance on the basis of 30-day all-cause readmission rates among patients with heart failure. *Circ Cardiovasc Qual Outcomes*. 2008;1(1):29–37. doi:10.1161/CIRCOUTCOMES.108.802686
24. Krumholz HM, Lin Z, Drye EE, et al. An administrative claims measure suitable for profiling hospital performance based on 30-day all-cause readmission rates among patients with acute myocardial infarction. *Circ Cardiovasc Qual Outcomes*. 2011;4(2):243–252. doi:10.1161/CIRCOUTCOMES.110.957498
25. Kanwal F, Tapper EB, Ho C, et al. Development of quality measures in cirrhosis by the practice metrics committee of the American association for the study of liver diseases. *Hepatology*. 2019;69(4):1787–1797. doi:10.1002/hep.30489
26. Kim DH, Glynn RJ, Avorn J, et al. Validation of a claims-based frailty index against physical performance and adverse health outcomes in the health and retirement study. *J Gerontol a Biol Sci Med Sci*. 2019;74(8):1271–1276. doi:10.1093/gerona/gly197
27. Kim DH, Patorno E, Pawar A, et al. Measuring frailty in administrative claims data: comparative performance of four claims-based frailty measures in the US medicare data. *J Gerontol a Biol Sci Med Sci*. 2020;75(6):1120–1125. doi:10.1093/gerona/glz224
28. Gautam N, Bessette L, Pawar A, et al. Updating international classification of diseases 9th revision to 10th revision of a claims-based frailty index. *J Gerontol a Biol Sci Med Sci*. 2021;76(7):1316–1317. doi:10.1093/gerona/glaa150
29. Gagne JJ, Glynn RJ, Avorn J, et al. A combined comorbidity score predicted mortality in elderly patients better than existing scores. *J Clin Epidemiol*. 2011;64(7):749–759. doi:10.1016/j.jclinepi.2010.10.004
30. Lip GY, Nieuwlaat R, Pisters R, et al. Refining clinical risk stratification for predicting stroke and thromboembolism in atrial fibrillation using a novel risk factor-based approach: the euro heart survey on atrial fibrillation. *Chest*. 2010;137(2):263–272. doi:10.1378/chest.09-1584