A Methodology to Generate Longitudinally Updated Acute-On-Chronic Liver Failure Prognostication Scores From Electronic Health Record Data

Jin Ge [1], Nader Najafi,[2] Wendi Zhao,[2] Ma Somsouk,[1] Margaret Fang,[2] and Jennifer C. Lai [1]

Queries of electronic health record (EHR) data repositories allow for automated data collection. These techniques have not been used in hepatology due to the inability to capture hepatic encephalopathy (HE) grades, which are inputs for acute-on-chronic liver failure (ACLF) models. Here, we describe a methodology to use EHR data to calculate rolling ACLF scores. We examined 239 patient admissions with end-stage liver disease from July 2014 to June 2019. We mapped EHR flowsheet data to determine HE grades and calculated two longitudinally updated ACLF scores. We validated HE grades and ACLF diagnoses by chart review and calculated sensitivity, specificity, and Cohen's kappa. Of 239 patient admissions analyzed, 37% were women, 46% were non-Hispanic white, median age was 60 years, and the median Model for End-Stage Liver Disease–Na score at admission was 25. Of the 239, 7% were diagnosed with ACLF as defined by the North American Consortium for the Study of End-Stage Liver Disease (NACSELD) diagnostic criteria at admission, 27% during the hospitalization, and 9% at discharge. Forty percent were diagnosed with ACLF by the European Association for the Study of the Liver– Chronic Liver Failure Consortium (CLIF-C) diagnostic criteria at admission, 51% during the hospitalization, and 34% at discharge. From the chart review of 51 admissions, we found sensitivities and specificities for any HE (grades 1-4) were 92%-97% and 76%-95%, respectively; for severe HE (grades 3-4), sensitivities and specificities were 100% and 78%-98%, respectively. Cohen's kappa between flowsheet and chart review of HE grades ranged from 0.55 to 0.72. Sensitivities and specificities for NACSELD-ACLF diagnoses were 75%-100% and 96%-100%, respectively; for CLIF-C-ACLF diagnoses, these were 91%-100% and 96-100%, respectively. We generated approximately 28 unique ACLF scores per patient per admission day. *Conclusion:* We developed an informatics-based methodology to calculate longitudinally updated ACLF scores. This opens new analytic potentials, such as big data methods, to develop electronic phenotypes for patients with ACLF. (*Hepatology Communications* 2021;5:1069-1080).

Electronic health records (EHRs) capture and generate vast amounts of granular clinical data through routine operations.[1] Structured Query Language (SQL) queries of associated clinical data repositories (CDRs) allow for automated generation of comprehensive laboratory, flowsheet, medical device, and medication administration reports for a cohort of patients.[2-4] Integration of these separate

Abbreviations: ACLF, acute-on-chronic liver failure; APASL, Asian Pacific Association for the Study of the Liver; CDR, clinical data repository; CI, confidence interval; CLIF, chronic liver failure; CLIF-C, Chronic Liver Failure Consortium; EASL, European Association for the Study of the Liver; EASL, European Association for the Study of the Liver; EHR, electronic health record; EPIC, EpicCare; ESLD, end-stage liver disease; FiO_2, oxygen fraction; FrAILT, Multi-Center Functional Assessment in Liver Transplantation Study; GCS, Glasgow Coma Score; HE, hepatic encephalopathy; IQR, interquartile range; LFI, Liver Frailty Index; MELD, Model for End-Stage Liver Disease; MELD-Na, Model for End-Stage Liver Disease–Sodium; NACSELD, North American Consortium for the Study of End-Stage Liver Disease; SQL, Structured Query Language; WHC, West Haven Criteria.

Received November 23, 2020; accepted January 24, 2021.

Additional Supporting Information may be found at onlinelibrary.wiley.com/doi/10.1002/hep4.1690/suppinfo.

Supported by the National Institute of Diabetes and Digestive and Kidney Diseases (5T32DK060414-18 to J.G.), the National Institute on Aging (R01AG059183/K23AG048337 to J.C.L.) and UCSF Liver Center Grant (P30 DK026743 to J.C.L.).

The funding agency played no role in the analysis of the data or the preparation of this manuscript.

data reports have the potential to survey patients in a longitudinal fashion during inpatient admission and construct electronic phenotypes to define subgroups of interest for further exploration.[5,6] Existing applications of SQL querying of data repositories in gastroenterology and hepatology, however, have been limited to searching International Classification of Diseases (ICD) 9/10 codes, identifying keywords in clinician documentation, and/or acquiring laboratory data.[7-9]

In hepatology research specifically, the adoption of informatic methods described above has been hindered by the inability to capture data to inform hepatic encephalopathy (HE) grades, which are often used as inputs into clinical prognostication models. EHR flowsheets, which contain structured and semistructured entries reflecting interprofessional assessments of mentation, functional status, and physical exam findings, represent a rich source of relevant clinical information.[3,4] These semistructured documentations of mentation have the potential to be mapped to describe HE, thereby enabling en masse automated data acquisition for clinical research in hepatology.

This becomes especially relevant in the study of acute-on-chronic liver failure (ACLF), which is defined as the acute decompensation of end-stage liver disease (ESLD) with extrahepatic organ failures and high short-term mortality.[10-16] ACLF is a heterogeneous and dynamic clinical syndrome with variable etiologies, triggers, and outcomes.[17] Reflecting the diversity of ACLF, several competing definitions and scoring systems currently exist, such as the North American Consortium for the Study of End-Stage Liver Disease (NACSELD) diagnostic criteria and NACSELD-ACLF score[10], the European Association for the Study of the Liver– Chronic Liver Failure (EASL-CLIF) diagnostic criteria and CLIF-Consortium-ACLF (CLIF-C-ACLF) score,[13] and the Asian Pacific Association for the Study of the Liver (APASL) ACLF Research Consortium definition and score.[16] Notably, the APASL ACLF definition differs in comparison to European (EASL-CLIF) and North American (NACSELD) definitions as it includes chronic liver diseases without cirrhosis and only considers precipitants that are intrahepatic in origin.[17,18]

Despite these differences in definitions of ACLF, the existing ACLF prognostication scores are generated in a cross-sectional manner at a specific point in time and remain limited in clinical utility due to inconsistent abilities to predict recovery and identify transplant candidates.[19] Moreover, the lack of consensus on a unified prognostication model implies that current methodologies for predictive modeling may be inadequate for this disease state. Longitudinally updated ACLF scores, therefore, may be able to improve predictive ability and better inform ACLF outcomes research as studies have shown score changes and trajectories have greater prognostic value.[20-23]

In this study, we describe a two-step methodology to generate longitudinally updated ACLF

View this article online at wileyonlinelibrary.com.
DOI 10.1002/hep4.1690

Potential conflict of interest: Nothing to report.

ARTICLE INFORMATION:

From the [1]Division of Gastroenterology and Hepatology, Department of Medicine, University of California San Francisco, San Francisco, CA, USA; [2]Division of Hospital Medicine, Department of Medicine, University of California San Francisco, San Francisco, CA, USA.

ADDRESS CORRESPONDENCE AND REPRINT REQUESTS TO:

Jennifer C. Lai, M.D., M.B.A.
Division of Gastroenterology and Hepatology
Department of Medicine
University of California San Francisco
513 Parnassus Avenue, UCSF Box 0538
San Francisco, CA 94143, USA
E-mail: jennifer.lai@ucsf.edu
Tel.: +1-415-476-6422

prognostication scores applicable to patients with ESLD in the United States (NACSELD-ACLF and CLIF-C-ACLF scores):

1. Calculate West Haven Criteria (WHC) grades of HE by mapping mentation and functional status descriptors in flowsheet reports and validating these mapped WHC grades through chart review.
2. Integrate mapped WHC grades with relationally linked reports of laboratory value, medical device data, and medication administration reports to generate longitudinally updated ACLF prognostication scores.

Materials and Methods

We examined all inpatient admissions in a 5-year period from July 1, 2014, through June 30, 2019, for the 1,918 patients enrolled in the Multi-Center Functional Assessment in Liver Transplantation (FrAILT) Study at a single academic medical center (University of California, San Francisco Medical Center) as of October 30, 2019. The FrAILT Study is a prospective longitudinal study of adult patients with ESLD awaiting liver transplantation and evaluated in the ambulatory care setting.[24] Hospital admissions for these patients were excluded if the admission took place after liver transplantation, was for a scheduled liver transplantation within 48 hours, hospital stays were ≤24 hours, or if transplantation took place before their enrollment in the FrAILT Study. If a patient had multiple hospitalizations, we analyzed the hospitalization immediately after the most recent Liver Frailty Index (LFI) assessment to isolate one admission per patient (patient admission). Of note, all patients who are listed for liver transplantation at our medical center are admitted to a dedicated multidisciplinary Liver Transplant Unit jointly attended by a hepatologist and a transplant surgeon for inpatient care. A flow diagram of the analyzed patient population from the FrAILT Study is shown in Fig. 1.

Baseline demographic and clinical data were extracted from the date of the latest outpatient LFI assessment. Race/ethnicity was classified into the following categories: white, black, Hispanic, Asian, Native American, or other. Etiologies of liver disease were categorized as: chronic hepatitis C, alcohol-associated, autoimmune/cholestatic, chronic hepatitis B, and other etiologies. Patients were considered to comorbid diagnoses of hypertension, diabetes, or coronary artery disease if they were reported in the EHR. The Institutional Review Board at the University of California, San Francisco, approved this study.

SQL DATA COLLECTION

For this cohort of 1,918 patients and eligible admissions, we queried the EpicCare (EPIC) (Epic Systems, Verona, WI) Clarity CDR hosted at the University of California, San Francisco Medical Center for reports containing data generated through routine care (Table 1). These reports were then linked relationally through two unique identifiers, medical record number and contact serial number, to each patient and admission.

WHC FOR ENCEPHALOPATHY

Structured predefined entries for nursing assessments of speech, cognition, orientation, and level of consciousness documented in flowsheet data reports were extracted from the CDR using SQL queries. These entries are recorded by nursing staff as part of their admission, shift-change, and per unit-determined nursing assessments, as mandated by the Joint Commission in the standards for Provision of Care, Treatment, and Services[25] and the American Nursing Association's Standards of Professional Nursing.[26,27] HE grades were mapped based on matching these entries with clinical descriptors used for grading HE in practice guidelines and in the HE Scoring Algorithm (Table 2).[28-31] Glasgow Coma Scores (GCSs) were also mapped with WHC grades, with GCS 15 mapping to grade 0, GCS 12-14 mapping to grade 2, GCS 4-11 mapping to grade 3, and GCS 3 mapping to grade 4.[30] If there were multiple data entries (e.g., entries for GCS, level of consciousness, and orientation) recorded at a given time or if the entries mapped to discrepant WHC grades, the maximum mapped WHC grade was used by default to maximize detection sensitivity. Structured data entries that did not fall under the above criteria, such as "Other (comment)," and unstructured data entries were excluded from mapping, thereby avoiding the need for natural language processing.

To validate the mapped WHC grades, we conducted blinded physician manual chart reviews of the

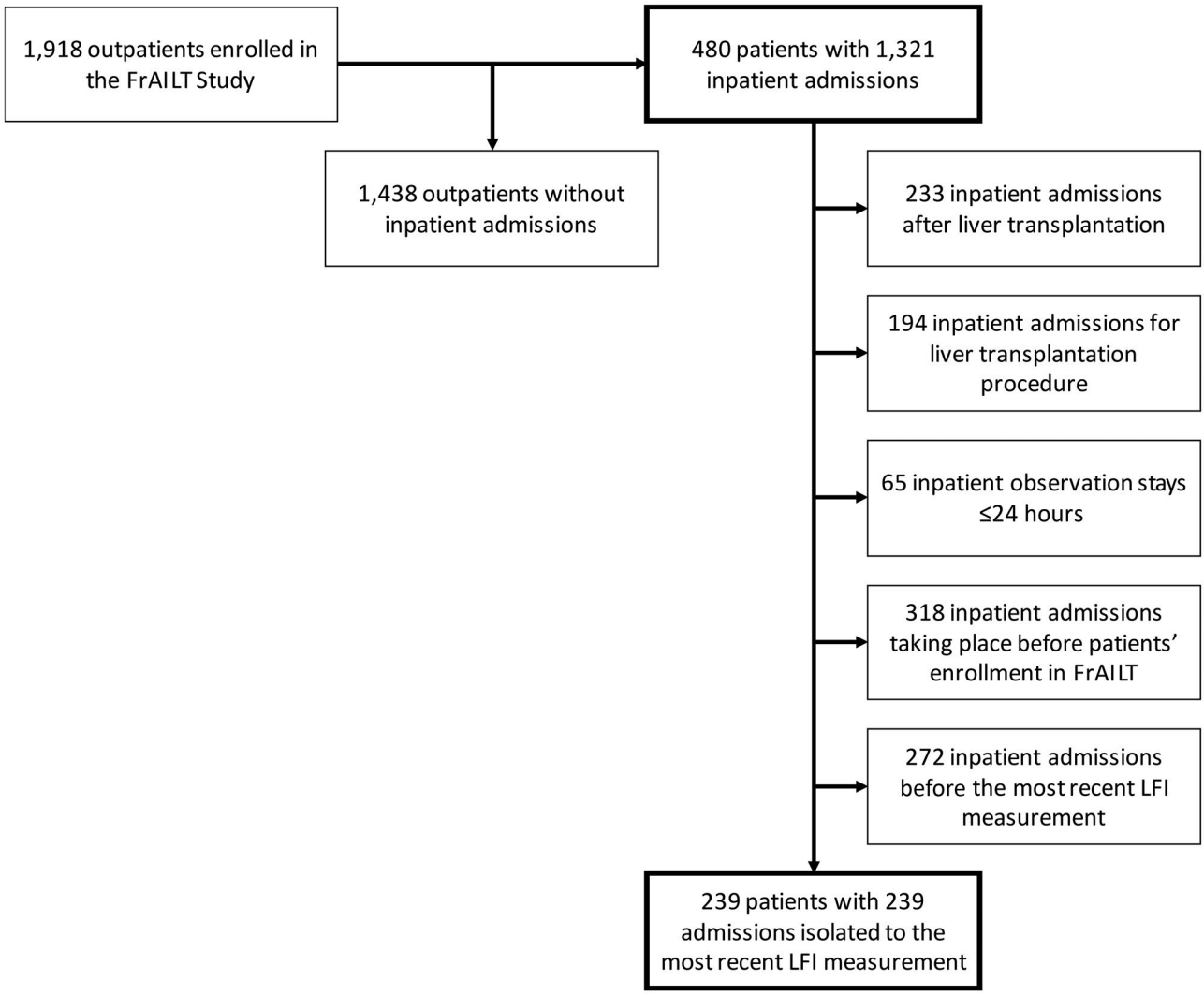

FIG. 1. Isolation of the 239 patient admissions analyzed in this study.

relevant sections (subjective findings, physical examination, and assessment and plan) of history and physical notes, progress notes, and discharge summaries of a random subset (20%) of patient admissions. The validators were blinded to the WHC HE grades determined by the flowsheet query to avoid contamination bias. Validation WHC grades were assigned based on descriptors in the subjective, physical examination, and/or assessment, and plan sections of the note were matched to clinical descriptions used in the American Association for the Study of Liver Diseases and EASL practice guidelines for management of HE.[28] For example, physical examination findings of "lethargy" on chart review would be considered consistent with WHC grade 2, while findings of "arousable to voice only" or "grossly confused" would be consistent with WHC grade 3 (Supporting Table S1). If there were discrepancies between physical examination findings among different members of the provider team, then we used findings based on a hierarchical read based on the level of training. For example, the attending physician's documented examination finding would be used over that of a resident physician. This blinded-physician review was conducted at 3 time points during each admission: at time of initial admission, during the hospitalization (defined as maximum value acquired during the admission), and at the time of discharge.

TABLE 1. DATA AND REPORT ELEMENTS GENERATED FROM SQL QUERIES OF CLARITY CDR

Clarity CDR Report Elements	
Flowsheet report	- Structured documentation of speech, cognition, orientation level, level of consciousness, and GCS in nursing flowsheets
	- Oxygen device, peripheral capillary oxygen saturation measurements, and blood pressures by peripheral and arterial measurements in vital signs flowsheets
laboratory data report	All laboratory (such as complete blood count, basic metabolic panel, liver function tests, coagulation parameters, albumin, and others) and bacterial culture (such as urine, central blood, peripheral blood, peritoneal fluid, sputum, and others) orders and results
Dialysis order report	All provider orders for hemodialysis, continuous venovenous hemofiltration, continuous venovenous hemodialysis, peritoneal dialysis, and ultrafiltration
Ventilator data report	Mechanical ventilation use (bilevel positive airway pressure or mechanical ventilator), ventilator mode (such as pressure support, assist control, and others), fraction of inspired oxygen (FiO_2), and partial pressure to fraction of oxygen ratio, if available
Vasopressor administration report	Medication administration records and times for all administrations of epinephrine, norepinephrine, phenylephrine, vasopressin, dopamine, and dobutamine

DETERMINATION OF OXYGEN FRACTION

Whenever available, the recorded oxygen fraction (FiO_2) in ventilator data or vital sign flowsheet reports were used. If such data were not available, then we estimated FiO_2 based on nasal cannula and high-flow nasal cannula flow rates, assuming closed-mouth breathing as validated in respiratory care literature (Supporting Table S2).[32-34] These recorded and estimated FiO_2 values were used to calculate oxygen saturation/FiO_2 and partial pressure to fraction of oxygen ratios, when appropriate.

ACLF DEFINITIONS AND PROGNOSTICATION SCORE CALCULATION

For each patient admission, we used the above data and mapped WHC grades to diagnose ACLF and calculated prognostication scores based on those published by the NACSELD-ACLF[10] and CLIF-C-ACLF.[13] The NACSELD-ACLF score (range 0-1) predicts the probability of 30-day survival in hospitalized patients (Supporting Table S3).[10] Model for End-Stage Liver Disease (MELD) and MELD-Na scores were also calculated, as described,[35,36] as inputs into the NACSELD-ACLF score. Similarly, the CLIF-C-ACLF score is a composite score (range, 0-100), with a score ≥70 predicting up to 100% mortality at 28 days (Supporting Table S3).[13] We did not use the APASL ACLF diagnostic criteria due to bacterial infection being the most common precipitant of ACLF in patients with cirrhosis in the United States, in contrast to those in the Asia-Pacific region.[19]

Using the mapped WHC grades from the methods above, we then generated automated diagnoses based on NACSELD and EASL-CLIF criteria at the 3 time points specified for validation of WHC grades of HE: at the time of initial admission, during the hospitalization (defined as maximum value acquired during the admission), and at the time of discharge. To validate the automated ACLF diagnoses, we conducted blinded-physician chart reviews of the relevant sections (physical examination, laboratory findings, and assessment and plan) of history and physical notes, progress notes, and discharge summaries of the same random subset (20%) of patient admissions validated in the WHC grade validation. NACSELD and EASL-CLIF ACLF diagnoses were confirmed based on descriptors in the physical examination findings, laboratory values, and diagnoses in the assessment and plan sections of the notes as matched to the relevant diagnostic criteria.[10,13]

STATISTICAL ANALYSES

Clinical characteristics and laboratory data for participants were summarized by medians and interquartile ranges (IQRs) for continuous variables or numbers and percentages for categorical variables. Comparisons among groups were performed using chi-square and Kruskal-Wallis tests, where appropriate. We calculated

TABLE 2. MAPPING OF FLOWSHEET DATA ON FOUR DOMAINS OF MENTAL STATUS AND GCS TO WHC FOR HE*

WHC[28-30]	Flowsheet Documentation of Mentation				GCS
	Speech	Cognition	Orientation Level	Level of Consciousness	
Grade 0. "No encephalopathy at all, no history of HE."	• Clear • Appropriate for developmental age • Uses written communication • Nods/gestures appropriately	• Appropriate judgment • Appropriate safety awareness • Appropriate attention/concentration • Appropriate for developmental age • Follows commands • No short-term memory loss	• Oriented ×4 (person, place, time, and situation) • Oriented to place • Oriented to time • Oriented to person • Oriented to situation • Appropriate for developmental age	• Alert • Awake • Responds to verbal	15
Grade 1. "Trivial lack of awareness, euphoria or anxiety, shortened attention span, impairment of addition or subtraction, altered sleep rhythm."	• Delayed responses	• Impulsive	• Disoriented to situation		
Grade 2. "Lethargy or apathy, disorientation for time, obvious personality change, inappropriate behavior, dyspraxia, asterixis."	• Word finding difficulty • Slurred	• Poor safety awareness • Poor judgement • Poor attention/concentration • Short-term memory loss	• Disoriented to person • Disoriented to time	• Lethargic	12-14
Grade 3. "Somnolence to semi stupor, responsive to stimuli, confused, gross disorientation, bizarre behavior."	• Incomprehensible • Expressive aphasia • Receptive aphasia	• Unable to follow commands	• Disoriented to place • Disoriented ×4	• Somnolent • Responds to pain only • Difficult to maintain arousal • Confused	4-11
Grade 4. "Coma."	• Global aphasia			• Obtunded • Unresponsive	3

*This classification scheme was based on guidance from descriptors in the WHC and criteria used in the HE Scoring Algorithm.[28-30]

sensitivities and specificities of the ability of the mapped WHC grades to detect any HE (WHC grades 1-4) and severe HE (WHC grades 3-4) uncovered by chart review. To rate interobserver agreement between mapped WHC grades and those acquired from chart review, we calculated Cohen's kappa coefficients and generated 95% confidence intervals (CIs) from bootstrapping with 1,000 replications.[37] Similarly, we also calculated sensitivities and specificities of the automated NACSELD-ACLF and EASL-CLIF ACLF diagnoses versus manual chart diagnoses. Two-sided $P < 0.05$ was considered statistically significant in all analyses. Analyses were performed using STATA statistical software, version 16.1 (StataCorp, College Station, TX).

Results

Of the 1,918 patients in the FrAILT Study at the University of California, San Francisco, 480 patients (25%) had 1,321 admissions during the 5-year study period. Of these 1,321 admissions, 233 occurred after liver transplantation, 194 were admissions for the liver transplantation surgical procedure, 65 hospital stays were ≤24 hours, and 318 transplantations took place before enrollment in the FrAILT Study. Of the 239 remaining patients meeting the inclusion criteria, we isolated one admission per patient for 239 patient admissions (Fig. 1).

BASELINE CHARACTERISTICS

Baseline characteristics of the 239 patients are presented in Table 3. Of these patients, 37% were women, 46% were non-Hispanic white, and median age at admission was 60 years (IQR, 53-65 years). The most common etiologies of cirrhosis were chronic hepatitis C (31%), alcohol-associated liver disease (24%), nonalcoholic fatty liver disease (21%), autoimmune/cholestatic diseases (10%), and chronic hepatitis B (5%). The median MELD score at admission was 21 (IQR, 15-29), and the median MELD-Na score at admission was 25 (IQR, 17-32).

TABLE 3. BASELINE CLINICAL AND DEMOGRAPHIC CHARACTERISTICS

Characteristic		FrAILT Cohort (n = 239)
Age in years at first admission (IQR)		60 (53-65)
Female (%)		88 (37)
Race/ethnicity (%)		
	White	111 (46)
	Black	9 (4)
	Hispanic	81 (34)
	Asian	23 (10)
	Native American	5 (2)
	Other	10 (4)
Etiology of liver disease (%)		
	Hepatitis C	75 (31)
	Alcoholic	57 (24)
	Nonalcoholic fatty	49 (21)
	AIH/PBC/PSC	25 (10)
	Hepatitis B	13 (5)
	Other etiologies	20 (8)
HCC (%)		71 (30)
Comorbidities (%)		
	Hypertension	112 (47)
	Diabetes	71 (30)
	Coronary artery disease	9 (4)
	Stroke	2 (1)
MELD at admission (IQR)		21 (15-29)
MELD-Na at admission (IQR)		25 (17-32)
NACSELD-ACLF at admission		17 (7)
	NACSELD-OF 0	179 (76)
	NACSELD-OF 1	40 (17)
	NACSELD-OF 2	16 (7)
	NACSELD-OF 3	1 (0.4)
	NACSELD-OF 4	0 (0)
CLIF-C-ACLF at admission		89 (40)
	CLIF-C-ACLF class 0	134 (60)
	CLIF-C-ACLF class 1	39 (17)
	CLIF-C-ACLF class 2	34 (15)
	CLIF-C-ACLF class 3	16 (7)
Total length of stay (IQR)		5 (2-9)

Abbreviations: AIH, autoimmune hepatitis; HCC, hepatocellular carcinoma; NACSELD-OF, North American Consortium for the Study of End-Stage Liver Disease–Organ Failures; PBC, primary biliary cirrhosis; PSC, primary sclerosing cholangitis.

VALIDATION OF WHC FOR HE

Given that validation of WHC grades using EHR had not previously been performed, we randomly selected 51 patient admissions (21%) to undergo manual chart review to validate the WHC grades mapped from flowsheet data. Sensitivities and specificities (along with 95% CIs) for the presence of any

HE and severe HE by comparing mapped WHC grades versus chart review are presented in Table 4. The sensitivities for the presence of any HE ranged from 92%-97%, while those for severe HE was 100% at the 3 time points queried (initial time of admission, during hospitalization, and time of discharge). Specificities for the presence of any HE ranged from 76%-95%, while those for severe HE ranged from 78%-98% at the 3 time points queried. Cohen's kappa coefficients for agreement between different WHC grades were 0.55 (95% CI, 0.33-0.74) at the time of admission, 0.64 (95% CI, 0.49-0.79) at the time of maximum value in the middle of the admission, and 0.72 (95% CI, 0.51-0.90) at the time of discharge.

ACLF DIAGNOSES

For the NACSELD-ACLF diagnostic criteria for ACLF, 17 patients (7%) were diagnosed at the time of initial admission, with 16 having two organ failures and 1 having three organ failures. The number of patients diagnosed with ACLF by the NACSELD criteria increased to 64 (27%) during the hospitalization, with 20 having two, 18 having three, and 26 having four organ failures. At the time of discharge, 21 patients (9%) still had ACLF diagnoses, with 7 having two, 6 having three, and 8 having four organ failures. In comparison to the manual chart review of the same 51 patient admissions selected for WHC validation, we found the sensitivities and specificities for ACLF diagnoses under the NACSELD criteria to be 75% and 96%, respectively, at the time of initial admission. The sensitivity and specificities increased to 100% and 97%, respectively, during the admission and then to 100% for both at the time of discharge (Table 5).

With respect to the EASL-CLIF ACLF diagnostic criteria, 89 patients (40%) were diagnosed at the time of initial admission, with 39 meeting grade 1, 34 grade 2, and 16 grade 3 criteria. The number of patients diagnosed with ACLF by the EASL-CLIF criteria increased to 114 (51%) during the hospitalization, with 28 meeting grade 1, 34 grade 2, and 52 grade 3 criteria. At the time of discharge, 76 (34%) still had ACLF diagnoses, with 35 meeting grade 1, 21 grade 2, and 19 grade 3 criteria. Similarly, when we compared these diagnoses to the manual chart review of the same 51 patient admissions selected for WHC validation, we found the sensitivities and specificities for ACLF diagnoses under the EASL-CLIF criteria to be 91% and 96%, respectively, at the time of initial admission. These figures increased to 100% for both during the admission and at the time of discharge (Table 5).

LONGITUDINAL ACLF PROGNOSTICATION SCORES

A total of 44,639 unique data points from the 239 patient admissions were available for ACLF prognostication score generation. This represented a median of 454 data points per admission (IQR, 194-704) and a median of 28 data points per admission day (IQR, 21-35). Using the data points generated from the relationally linked databases, we were able to calculate approximately hourly updated NACSELD-ACLF and CLIF-C-ACLF scores. A representative example of one patient's hospitalization course and

TABLE 4. SENSITIVITY AND SPECIFICITY OF MAPPED WHC VERSUS CHART REVIEW

	True Positive	False Negative	False Positive	True Negative	Sensitivity	95% CI	Specificity	95% CI
Any HE (WHC 1-4) at admission	12	1	9	29	0.92	0.64-1.00	0.76	0.60-0.89
Any HE (WHC 1-4) during hospitalization	31	1	3	16	0.97	0.84-1.00	0.84	0.60-0.97
Any HE (WHC 1-4) at discharge	11	1	2	37	0.92	0.61-1.00	0.95	0.83-0.99
Severe HE (WHC 3-4) at admission	4	0	6	41	1.00	0.40-1.00	0.87	0.74-0.95
Severe HE (WHC 3-4) during hospitalization	15	0	8	28	1.00	0.78-1.00	0.78	0.61-0.90
Severe HE (WHC 3-4) at discharge	5	0	1	45	1.00	0.48-1.00	0.98	0.89-1.00

TABLE 5. SENSITIVITY AND SPECIFICITY OF CALCULATED NACSELD-ACLF AND CLIF-C-ACLF DIAGNOSES VERSUS CHART REVIEW

Diagnosis	True Positive	False Negative	False Positive	True Negative	Sensitivity	95% CI	Specificity	95% CI
Dx of NACSELD-ACLF at admission	3	1	2	45	0.75	0.19-0.99	0.96	0.86-1.00
Dx of NACSELD-ACLF during hospitalization	15	0	1	36	1.00	0.78-1.00	0.97	0.86-1.00
Dx of NACSELD-ACLF at discharge	4	0	0	47	1.00	0.40-1.00	1.00	0.93-1.00
Dx of CLIF-C-ACLF at admission	21	2	1	27	0.91	0.72-0.99	0.96	0.82-1.00
Dx of CLIF-C-ACLF during hospitalization	27	0	0	24	1.00	0.87-1.00	1.00	0.86-1.00
Dx of CLIF-C-ACLF at discharge	19	0	0	32	1.00	0.82-1.00	1.00	0.89-1.00

Abbreviation: Dx, diagnosis.

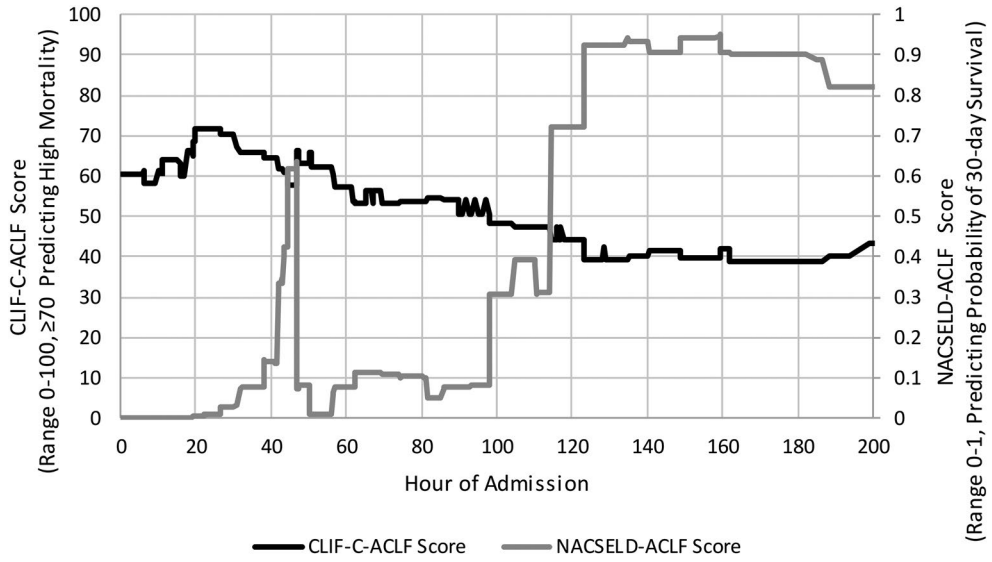

FIG. 2. Representative sample of calculated longitudinal NACSELD-ACLF and CLIF-C-ACLF scores.

the corresponding NACSELD-ACLF and CLIF-C-ACLF scores are shown in Fig. 2.

Discussion

In this study, we validated an informatics-based method for capturing routine clinical care data from the EHR to calculate longitudinally updated ACLF prognostication scores in patients with ESLD hospitalized for ACLF. We extracted and mapped nursing assessment flowsheet data, which are captured at regular intervals, to WHC grades for HE and integrated these mapped values to more traditional data (laboratory values, device data, and medication administration records) to diagnose ACLF and calculate longitudinally updated prognostic scores under two definitions (NACSELD-ACLF and CLIF-C-ACLF scores; Fig. 2). Our method demonstrated high sensitivity in detecting any or severe HE throughout the

hospitalization and moderate sensitivity for diagnosing ACLF at admission that rapidly improved during the hospitalization.

While the techniques of using SQL queries of CDRs have been demonstrated,[1] existing applications have been limited.[7-9] Our method is novel in that it integrates the following multiple sources of EHR data: laboratory data, medication administration records, provider orders, ventilator device data, and flowsheet data. We made extensive use of flowsheets, which contain interprofessional (particularly nursing) assessments of mentation, functional status, and examination, that contained structured data entries that mapped to WHC grades based on previously validated instruments. Flowsheet data comprise one third of all recorded data in CDRs and have been historically underused in clinical research.[3,4] Indeed, the linchpins in our methodology were mapping of flowsheet entries to appropriate WHC grades and of supplemental oxygenation flow rates to estimated FiO_2.

Validation of WHC grades generated from flowsheet data proved to have high sensitivity (92%-95% for any HE and 100% for severe HE) versus clinicians' documentation. Of note, HE grading has historically been difficult due to subjective assessments with poor to moderate interrater reliability.[28,38-41] This appears to be the rationale behind the use of overt HE, which has interrater reliability for diagnosing brain failure in both NACSELD and EASL-CLIF definitions.[38] Our calculations of Cohen's kappa coefficient indicated moderate agreement between WHC grades mapped from flowsheet data and those rated by clinicians on retrospective chart review. While we found that Cohen's kappa coefficients were lower at the beginning of the admission (0.55 vs. 0.72 at the time of discharge), this was thought to be due to lack of synchronization between clinician and nursing documentation, which gradually improved during the hospitalizations analyzed. The kappa coefficients from our study (0.55-0.72) are within range of those reported for other methodologies for differentiating WHC grades.[30,42,43] While the sensitivities and specificities of our methodology for ACLF diagnosis (based on NACSELD and EASL-CLIF criteria) were imperfect at the time of initial admission, they rapidly improved to 100% as more information was generated and gathered throughout the admission.

We acknowledge the following limitations to our study. The first is that our methodology was developed at a single center and on a specific implementation of the EPIC EHR system without validation at another site. As such, our methodology and findings may not be readily generalizable to other centers (such as nontransplant centers) with non-EPIC EHR systems. While specific structured data elements may differ in EHR flowsheets between institutions, the nursing documentation of neurologic and mentation assessments (level of consciousness, orientation, and cognition) used in our methodology are mandated nationally by the Joint Commission in the standards for Provision of Care, Treatment, and Services[25] and the American Nursing Association's Standards of Professional Nursing[26,27] and locally by the California Nursing Practice Act.[44] Review of charting and nursing standards of care documents at hospitals affiliated with our institution (academic medical center, county safety-net hospital, and Veterans Affairs hospital) showed consistency in the detail of documentation in these assessments. Given prescribed standards for nursing assessments and documentations by national and local accreditation bodies, we anticipate that the data recorded in flowsheets are likely to be similar at other institutions and across EHR platforms. In our future works, we intend to replicate and validate this methodology at other institutions (using EPIC-based data available from other University of California Health medical centers) and on other EHR platforms, such as Cerner.

Even with this first limitation, the major advantage of our methodology is that it uses existing documentation practices and charting infrastructure to discern different gradations of HE. While the exact execution for this methodology (e.g., SQL extraction code) will differ at another institution, the general strategy of mapping and extracting flowsheet data remains the same. In our experience, while SQL code written for our institution's CDR does not often work out-of-the-box against that of another institution, the differences are generally correctable and reconcilable. Moreover, ongoing efforts in reconciling flowsheet data to clinically focused information models, such as with regards to pain[45] and genitourinary domains,[46] using common data codes (e.g., Logical Observation Identifier, Name, and Codes and Systematized Nomenclature of Medicine)[47,48] will allow for greater consistency across institutions. The movement toward standard data models, such as the Observational Medical Outcomes Partnership, which is used at more than 150

institutions and includes elements for flowsheet data, will also greatly improve future interoperability for the development of robust multicenter collaborations.[49]

Second, the patient population evaluated in this study was highly selected because we only considered patients with ESLD who were enrolled in the FrAILT Study (evaluated for transplantation). This high degree of selection, however, was by design to validate our methodology in a controlled cohort. The next logical extension of this methodology is for implementations in larger and nonspecific cohorts, such as all patients with ICD-9/10 discharge diagnoses of cirrhosis.

Third, we relied on blinded-physician manual chart review of subjective findings, physical examination, and assessment and plans documented as the standard for validation. This is a limitation due to the retrospective nature of our study as many treating clinicians do not necessarily perform or document West Haven assessments in routine clinical care.

Last, given that the University of California, San Francisco Medical Center is a tertiary referral center, many of the admissions evaluated in our study were transferred to our medical center. Patient admissions in our sample, thus, may not reflect the initial clinical course. Moreover, we suspect that initial delays in documentation or clinician order placement (such as entering orders for dialysis) for transfer admissions contributed to the relatively poor sensitivity of our method for diagnosis of NACSELD ACLF at admission compared to chart review. As expected, the accuracy and precision of our methodology increased through the length of stay as more data were integrated.

Despite these limitations, this study serves as a proof of concept for a clinical informatics-based methodology to generate longitudinally updated ACLF prognostication scores, which can better reflect the dynamic clinical course of these patients. Pilot demonstration of the validity of this methodology to extract accurate data in this population opens new analytic potentials, such as the application of big data methods that leverage the rich data from EHR platforms and CDR configurations to enhance investigation of predictors of outcomes in this dynamic population.

REFERENCES

1) Atreja A, Achkar J-P, Jain AK, Harris CM, Lashner BA. Using technology to promote gastrointestinal outcomes research: a case for electronic health records. Am J Gastroenterol 2008;103:2171-2178.
2) Milinovich A, Kattan MW. Extracting and utilizing electronic health data from Epic for research. Ann Transl Med 2018;6:42.
3) Waitman LR, Warren JJ, Manos EL, Connolly DW. Expressing observations from electronic medical record flowsheets in an i2b2 based clinical data repository to support research and quality improvement. AMIA Annu Symp Proc 2011;2011:1454-1463.
4) Johnson SG, Byrne MD, Christie B, Delaney CW, LaFlamme A, Park JI, et al. Modeling flowsheet data for clinical research. AMIA Jt Summits Transl Sci Proc 2015;2015:77-81.
5) Overby CL, Pathak J, Gottesman O, Haerian K, Perotte A, Murphy S, et al. A collaborative approach to developing an electronic health record phenotyping algorithm for drug-induced liver injury. J Am Med Inform Assoc 2013;20:e243-e252.
6) Newton KM, Peissig PL, Kho AN, Bielinski SJ, Berg RL, Choudhary V, et al. Validation of electronic medical record-based phenotyping algorithms: results and lessons learned from the eMERGE network. J Am Med Inform Assoc 2013;20:e147-e154.
7) Heidemann L, Law J, Fontana RJ. A text searching tool to identify patients with idiosyncratic drug-induced liver injury. Dig Dis Sci 2017;62:615-625.
8) Rudrapatna VA, Glicksberg BS, Avila P, Harding-Theobald E, Wang C, Butte AJ. Accuracy of medical billing data against the electronic health record in the measurement of colorectal cancer screening rates. BMJ Open Qual 2020;9:e000856.
9) Anderson AJM, Click B, Ramos-Rivers C, Babichenko D, Koutroubakis IE, Hartman DJ, et al. Development of an inflammatory bowel disease research registry derived from observational electronic health record data for comprehensive clinical phenotyping. Dig Dis Sci 2016;61:3236-3245.
10) O'Leary JG, Reddy KR, Garcia-Tsao G, Biggins SW, Wong F, Fallon MB, et al. NACSELD acute-on-chronic liver failure (NACSELD-ACLF) score predicts 30-day survival in hospitalized patients with cirrhosis. Hepatology 2018;67:2367-2374.
11) Moreau R, Jalan R, Gines P, Pavesi M, Angeli P, Cordoba J, et al.; CANONIC Study Investigators of the EASL–CLIF Consortium. Acute-on-chronic liver failure is a distinct syndrome that develops in patients with acute decompensation of cirrhosis. Gastroenterology 2013;144:1426-1437, 1437.e1-e.9.
12) Gustot T, Fernandez J, Garcia E, Morando F, Caraceni P, Alessandria C, et al.; CANONIC Study Investigators of the EASL-CLIF Consortium. Clinical course of acute-on-chronic liver failure syndrome and effects on prognosis. Hepatology 2015;62:243-252.
13) Jalan R, Saliba F, Pavesi M, Amoros A, Moreau R, Ginès P, et al.; CANONIC study investigators of the EASL-CLIF Consortium. Development and validation of a prognostic score to predict mortality in patients with acute-on-chronic liver failure. J Hepatol 2014;61:1038-1047.
14) Bajaj JS, O'Leary JG, Reddy KR, Wong F, Biggins SW, Patton H, et al.; North American Consortium For The Study Of End-Stage Liver Disease (NACSELD). Survival in infection-related acute-on-chronic liver failure is defined by extrahepatic organ failures. Hepatology 2014;60:250-256.
15) Sarin SK, Kumar A, Almeida JA, Chawla YK, Fan ST, Garg H, et al. Acute-on-chronic liver failure: consensus recommendations of the Asian Pacific Association for the study of the liver (APASL). Hepatol Int 2009;3:269-282.
16) Sarin SK, Choudhury A, Sharma MK, Maiwall R, Al Mahtab M, Rahman S, et al.; APASL ACLF Research Consortium (AARC) for APASL ACLF working Party. Acute-on-chronic liver failure: consensus recommendations of the Asian Pacific association for the study of the liver (APASL): an update. Hepatol Int 2019;13:353-390.
17) Hernaez R, Solà E, Moreau R, Ginès P. Acute-on-chronic liver failure: an update. Gut 2017;66:541-553.

18) Kim TY, Song DS, Kim HY, Sinn DH, Yoon EL, Kim CW, et al.; Korean Acute-on-Chronic Liver Failure Study Group. Characteristics and discrepancies in acute-on-chronic liver failure: need for a unified definition. PLoS One 2016;11:e0146745.
19) Bajaj JS, Moreau R, Kamath PS, Vargas HE, Arroyo V, Reddy KR, et al. Acute-on-chronic liver failure: getting ready for prime time? Hepatology 2018;68:1621-1632.
20) Choudhury A, Jindal A, Maiwall R, Sharma MK, Sharma BC, Pamecha V, et al.; APASL ACLF Working Party. Liver failure determines the outcome in patients of acute-on-chronic liver failure (ACLF): comparison of APASL ACLF research consortium (AARC) and CLIF-SOFA models. Hepatol Int 2017;11:461-471.
21) Rathi S, Taneja S, Duseja A, Gautam V, Chawla Y, Dhiman RK. Dynamic assessment is superior to baseline assessment in prognostication of patients with acute on chronic liver failure. J Hepatol 2018;68:S240-S241.
22) Mahmud N, Sundaram V, Kaplan DE, Taddei TH, Goldberg DS. Grade 1 acute on chronic liver failure is a predictor for subsequent grade 3 failure. Hepatology 2020;72:230-239.
23) Sundaram V, Jalan R, Wu T, Volk ML, Asrani SK, Klein AS, et al. Factors associated with survival of patients with severe acute-on-chronic liver failure before and after liver transplantation. Gastroenterology 2019;156:1381-1391.e3.
24) Lai JC, Rahimi RS, Verna EC, Kappus MR, Dunn MA, McAdams-DeMarco M, et al. Frailty associated with waitlist mortality independent of ascites and hepatic encephalopathy in a multicenter study. Gastroenterology 2019;156:1675-1682.
25) The Joint Commission. Provision of care, treatment, and services. The Joint Commission Comprehensive Accreditation Manual for Hospitals E-dition. 2020.
26) American Nurses Association. Nursing: Scope and Standards of Practice. 2nd ed. Silver Spring, MD: American Nurses Association; 2010.
27) McBride S, Tietze M. Nursing Informatics for the Advanced Practice Nurse, Second Edition: Patient Safety, Quality, Outcomes, and Interprofessionalism. New York, NY: Springer Publishing Company; 2018.
28) Vilstrup H, Amodio P, Bajaj J, Cordoba J, Ferenci P, Mullen KD, et al. Hepatic encephalopathy in chronic liver disease: 2014 Practice Guideline by the American Association for the Study of Liver Diseases and the European Association for the Study of the Liver. Hepatology 2014;60:715-735.
29) Hassanein TI, Hilsabeck RC, Perry W. Introduction to the hepatic encephalopathy scoring algorithm (HESA). Dig Dis Sci 2008;53:529-538.
30) Hassanein T, Blei AT, Perry W, Hilsabeck R, Stange J, Larsen FS, et al. Performance of the hepatic encephalopathy scoring algorithm in a clinical trial of patients with cirrhosis and severe hepatic encephalopathy. Am J Gastroenterol 2009;104:1392-1400.
31) Conn HO, Leevy CM, Vlahcevic ZR, Rodgers JB, Maddrey WC, Seeff L, et al. Comparison of lactulose and neomycin in the treatment of chronic portal-systemic encephalopathy. A double blind controlled trial. Gastroenterology 1977;72:573-583.
32) Wettstein RB, Shelledy DC, Peters JI. Delivered oxygen concentrations using low-flow and high-flow nasal cannulas. Respir Care 2005;50:604-609.
33) Ward JJ. High-flow oxygen administration by nasal cannula for adult and perinatal patients. Respir Care 2013;58:98-122.
34) Chikata Y, Onodera M, Oto J, Nishimura M. FIO2 in an adult model simulating high-flow nasal cannula therapy. Respir Care 2017;62:193-198.
35) Malinchoc M, Kamath PS, Gordon FD, Peine CJ, Rank J, ter Borg PC. A model to predict poor survival in patients undergoing transjugular intrahepatic portosystemic shunts. Hepatology 2000;31:864-871.
36) Kim WR, Biggins SW, Kremers WK, Wiesner RH, Kamath PS, Benson JT, et al. Hyponatremia and mortality among patients on the liver-transplant waiting list. N Engl J Med 2008;359:1018-1026.
37) Reichenheim ME. Confidence intervals for the kappa statistic. Stata J 2004;4:421-428.
38) Montagnese S, Amodio P, Morgan MY. Methods for diagnosing hepatic encephalopathy in patients with cirrhosis: a multidimensional approach. Metab Brain Dis 2004;19:281-312.
39) Bajaj JS, Cordoba J, Mullen KD, Amodio P, Shawcross DL, Butterworth RF, et al.; International Society for Hepatic Encephalopathy and Nitrogen Metabolism (ISHEN). Review article: the design of clinical trials in hepatic encephalopathy–an International Society for Hepatic Encephalopathy and Nitrogen Metabolism (ISHEN) consensus statement. Aliment Pharmacol Ther 2011;33:739-747.
40) Kircheis G, Fleig WE, Görtelmeyer R, Grafe S, Häussinger D. Assessment of low-grade hepatic encephalopathy: a critical analysis. J Hepatol 2007;47:642-650.
41) Hadjihambi A, Arias N, Sheikh M, Jalan R. Hepatic encephalopathy: a critical current review. Hepatol Int 2018;12(Suppl. 1): 135-147.
42) Edwin N, Peter JV, John G, Eapen CE, Graham PL. Relationship between clock and star drawing and the degree of hepatic encephalopathy. Postgrad Med J 2011;87:605-611.
43) Amodio P, Campagna F, Olianas S, Iannizzi P, Mapelli D, Penzo M, et al. Detection of minimal hepatic encephalopathy: normalization and optimization of the Psychometric Hepatic Encephalopathy Score. A neuropsychological and quantified EEG study. J Hepatol 2008;49:346-353.
44) California Board of Registered Nursing. Nursing Practice Act. https://www.rn.ca.gov/practice/npa.shtml. Published 2020. Accessed December 2020.
45) Westra BL, Johnson SG, Ali S, Bavuso KM, Cruz CA, Collins S, et al. Validation and refinement of a pain information model from EHR flowsheet data. Appl Clin Inform 2018;9:185-198.
46) Westra BL, Lytle KS, Whittenburg L, Adams M, Ali S, Furukawa M, et al. A refined methodology for validation of information models derived from flowsheet data and applied to a genitourinary case. J Am Med Inform Assoc 2020;27:1732-1740.
47) Matney S, Bakken S, Huff SM. Representing nursing assessments in clinical information systems using the logical observation identifiers, names, and codes database. J Biomed Inform 2003;36:287-293.
48) Matney SA, Settergren TT, Carrington JM, Richesson RL, Sheide A, Westra BL. Standardizing physiologic assessment data to enable big data analytics. West J Nurs Res 2017;39:63-77.
49) Stang PE, Ryan PB, Racoosin JA, Overhage JM, Hartzema AG, Reich C, et al. Advancing the science for active surveillance: rationale and design for the Observational Medical Outcomes Partnership. Ann Intern Med 2010;153:600-606.

Supporting Information

Additional Supporting Information may be found at onlinelibrary.wiley.com/doi/10.1002/hep4.1690/suppinfo.