Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

J Hepatol. Author manuscript; available in PMC 2023 May 01.

Published in final edited form as:

J Hepatol. 2022 May ; 76(5): 1122–1126. doi:10.1016/j.jhep.2021.12.040.

Discordance in categorization of acute-on-chronic liver failure in the United Network for Organ Sharing database

Brian P. Lee, MD MAS[*,1], **Giuseppe Cullaro, MD MAS**[*,2], **Aidan Vosooghi, BA**[1], **Frederick Yao, BA**[2], **Sarjukumar Panchal, MD**[3], **David S. Goldberg, MD MSCE**[4], **Norah A. Terrault, MD MPH**[1], **Nadim Mahmud, MD MSCE**[3]

[1]University of Southern California Keck School of Medicine, Los Angeles, CA

[2]University of California San Francisco, San Francisco, CA

[3]University of Pennsylvania Perelman School of Medicine, Philadelphia, PA

[4]University of Miami Leonard M. Miller School of Medicine, Miami, FL

Abstract

Background & Aims: Studies regarding acute-on-chronic liver failure (ACLF) among liver transplant (LT) candidates from the United Network for Organ Sharing (UNOS) database are being used to inform LT policy changes worldwide. We assessed the validity of identifying ACLF in UNOS.

Methods: We performed stratified random sampling among 3 US LT centers between 2013–2019 to obtain a representative patient sample across ACLF grades. We compared the concordance of ACLF classification by UNOS vs. blinded manual chart review, according to EASL-CLIF.

Results: Among 481 sampled LT registrants, 250 (52%) had no ACLF, 75 (16%) had ACLF grade 1, 79 (16%) had ACLF Grade 2, and 77 (16%) had ACLF Grade 3 per UNOS categorization. Concordance of ACLF grade by UNOS vs. chart review was: 72%, 64%, 56%,

Correspondence Address reprint requests to Brian Lee at 2250 Alcazar St, Room 135G, Los Angeles CA 90089, brian.lee@med.usc.edu.

*Co-first authors

Author Contributions

- Brian Lee: study concept and design, acquisition of data, analysis and interpretation of data, drafting of manuscript, revision of manuscript, statistical analysis, study supervision
- Giuseppe Cullaro: study concept and design, acquisition of data, analysis and interpretation of data, revision of manuscrip
- Aidan Vosooghi: acquisition of data, analysis and interpretation of data, revision of manuscript
- Frederick Yao: acquisition of data, analysis and interpretation of data, revision of manuscript
- Sarjukumar Panchal: acquisition of data, analysis and interpretation of data, revision of manuscript
- David S. Goldberg: study concept and design, analysis and interpretation of data, revision of manuscript
- Norah Terrault: study concept and design, analysis and interpretation of data, revision of manuscript, study supervision
- Nadim Mahmud: study concept and design, acquisition of data, analysis and interpretation of data, revision of manuscript

Disclosures: The authors have no potential conflicts of interest to disclose.

Data Availability Statement: The authors confirm that the data supporting the findings of this study are available within the article and its supplementary material.

Publisher's Disclaimer: This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. 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.

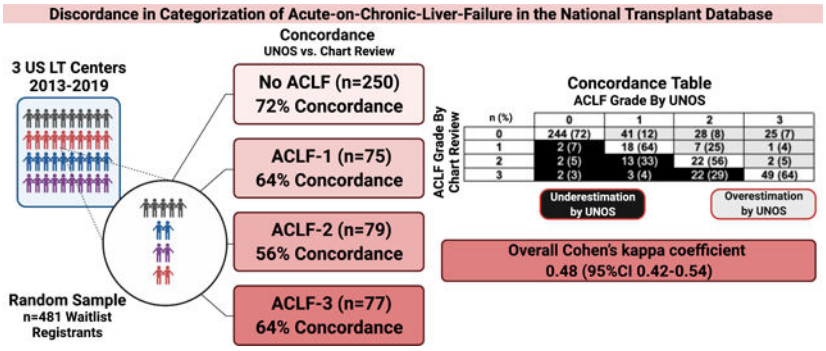
Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and 64% for no ACLF, Grade 1, Grade 2, and Grade 3, respectively, with overall Cohen's kappa coefficient 0.48 (95%CI 0.42–0.54). Absence of acute decompensation was the most common reason for overestimation, and discordant brain and respiratory failure categorization were the most common reasons for underestimation of ACLF by UNOS.

Conclusions: In this retrospective multi-center study, ACLF categorization by UNOS showed weak agreement with manual chart review. These findings are informative for ongoing allocation policy discussions, highlight the importance of prospective studies regarding ACLF in LT and should encourage UNOS reform.

Lay summary: Acute-on-chronic-liver-failure (ACLF) is a specific and common form of liver failure with high death rates. Studies have been published using the United States transplant registry (UNOS) to identify and describe outcomes of transplant candidates and recipients with ACLF, and these data are driving policy changes for transplant allocation around the world, but nobody has shown whether these data are reliable. We found that UNOS was not categorizing ACLF in concordance or accurately when compared to chart review, which shows the need for UNOS reform and non-UNOS studies to appropriately inform policies in transplant with ACLF.

Graphical Abstract



Keywords

UNOS; ACLF; concordance; misclassification; bias

Acute-on-chronic-liver failure (ACLF) is a distinct clinical entity in which decompensation of chronic liver disease occurs rapidly in the presence of extrahepatic organ failures.[1] ACLF has been associated with poor outcomes, independent of other prognostic factors (e.g. Model for End-Stage Liver Disease Sodium [MELD-Na] score) among hospitalized patients with cirrhosis.[2] ACLF, as a distinct prognostic factor, has led to recent calls to incorporate and prioritize ACLF for liver transplantation (LT) in the United States and Europe.[3]

LT candidates represent a unique population with end-stage liver disease with specific research outcomes (e.g. waitlist mortality), which typically require dedicated studies to affect allocation policy. While some small European studies have described waitlist and post-LT outcomes among waitlist registrants with ACLF,[3,4] discussions regarding LT allocation policy have largely been based on results from the United Network for Organ Sharing (UNOS) database. Indeed, several high-profile UNOS studies focused on waitlist

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

outcomes have described ACLF in separate studies as independently predicting waitlist mortality[3,5], having higher waitlist mortality than status 1a patients,[5] having interactions with obesity to predict waitlist mortality,[6] and having different trajectories of waitlist outcomes by grade of ACLF.[7] Countries such as the United Kingdom and Spain have made recommendations to prioritize ACLF in organ allocation, specifically citing these UNOS studies.[3] While these studies have prompted transplant policy discussions (e.g. granting ACLF exception points on LT waitlist) across the United States and Europe, the validity of identifying and correctly classifying ACLF in UNOS has yet to be studied.

While UNOS has many strengths including full national representation and reliable ascertainment of survival outcomes, we hypothesized that its use as a research database for ACLF may be flawed as not all variables within ACLF criteria exist in UNOS, and certain UNOS variables have been shown to be prone to misclassification. Specifically, UNOS does not capture the acuity of liver decompensation (i.e. time from decompensation to waitlist registration), has no existing variables to indicate gastrointestinal bleeding, bacterial infection, or data regarding oxygenation status beyond mechanical ventilation, and liver decompensations (e.g. ascites and encephalopathy) are provided with unclear accuracy.

To inform ongoing policy discussions and how to better advance ACLF research in future studies, we sought to assess: i) metrics of concordance in identifying ACLF in UNOS (vs. manual chart review); and ii) modifiable factors contributing to potential discordance.

METHODS

Study Population

Three LT centers (University of Southern California, University of California San Francisco, University of Pennsylvania) participated in this study. We performed stratified random sampling through sequential steps to obtain a representative patient sample. First, we identified all waitlist registrants with age≥18 at each site between 2013–2019 (to coincide with the institution of Share-35 policy and available follow-up at the time of this analysis). We excluded patients with acute liver failure, multiorgan transplants except simultaneous liver-kidney recipients, re-LT recipients, living donor recipients, HIV, and MELD-Na score exceptions, as they represent special circumstances that may affect waitlist outcomes not related to our research question. Second, we classified all patients by ACLF grade (no ACLF, Grade 1, Grade 2, Grade 3) according to the EASL-CLIF based algorithm described in previously published ACLF UNOS studies[7] (Supplemental Table 1). Finally, we randomly sampled 250 patients with no ACLF, 75 with ACLF Grade 1, 79 with ACLF Grade 2, and 77 with Grade 3 by UNOS classification, to ensure sufficient sample size for each ACLF grade for concordance metrics and regression analyses.

Classification of ACLF by Chart Review

After patients were sampled, investigators at each site (A.V. at USC, F.Y. at UCSF, and S.P. at UPenn) who were blinded to the patients' UNOS classification of ACLF (i.e. without access to UNOS data) their electronic medical record to complete a standardized data collection form. Formal metrics of data collection accuracy were not calculated, but the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

data collection form was piloted by all authors to ensure any possible misunderstandings were resolved prior to formal data collection. Listing demographics including sex, age, race, primary insurance status were captured, in addition to clinical characteristics at listing (primary etiology of liver disease, diabetes, MELD-Na score, hospitalization status, intensive care unit status). Alcohol-associated hepatitis (AH) was defined as meeting definite or probable AH by NIAAA consensus criteria[8]. Characteristics of ACLF at listing were captured and scored by organ failure according to EASL-CLIF criteria[1] (Supplemental Table 1). Encephalopathy was classified by West-Haven Criteria[1]. Reason for mechanical ventilation was categorized as encephalopathy/mental status, respiratory failure, or other. Acute liver decompensation was defined[1] as acute development of large ascites, encephalopathy, gastrointestinal bleeding, or bacterial infection within the preceding 4 weeks from listing. Patients with history of prior ascites or encephalopathy could be classified as developing new ascites or new encephalopathy if they met conditions specified by EASL-CLIF criteria[1]. Specifically, patients with history of prior ascites could qualify for new ascites with new development of grade 2 to 3 ascites (by International Ascites Club Classification).[1] Patients with chronic refractory ascites admitted to the hospital frequently for therapeutic paracentesis are not included in this definition.[1] Patients with history of prior encephalopathy could qualify for new encephalopathy with acute development of a change in mental status in a patient with previous normal consciousness and no evidence of an acute neurologic disease.[1] Details regarding categorization of each organ failure by chart review vs. UNOS are summarized in Supplemental Table 1. Infections were defined as those meeting specific criteria summarized in Supplemental Table 2.

Statistical Analysis

Demographic and clinical characteristics were described using means (SDs), medians [interquartile ranges (IQR)], and proportions as appropriate. Categorical variables were compared using the chi-square test.

Concordance and Accuracy Metrics—The primary aim, concordance of ACLF classification by UNOS, was compared to manual chart review. Discordant ACLF categorization was defined as any patient that did not have the same ACLF categorization in UNOS vs. chart review. To determine concordance across all ACLF categories, we also computed an unweighted Cohen's kappa using the kappaetc STATA package (https://ideas.repec.org/c/boc/bocode/s458283.html). This metric accounts for the possibility of agreement occurring by chance. A kappa value of ≥0.8 was regarded to indicate strong agreement, ≥0.6 to <0.8 moderate agreement, and ≥0.4 to <0.6 weak agreement.[9]

As a secondary aim, we also calculated accuracy metrics of ACLF classification by UNOS (vs. by chart review), calculating negative predictive value, positive predictive value, sensitivity, specificity, and receiver operating characteristic (ROC) area under the curve (AUC), with 95% confidence intervals, using the diagt STATA package (https://ideas.repec.org/c/boc/bocode/s423401.html). Factors associated with discordant classification were assessed by logistic regression, evaluating characteristics in Table 1 as potential factors associated with a discordant ACLF categorization as the outcome. In a fully-adjusted model, etiology of liver disease was collapsed for sample size reasons.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Intensive care unit status and alcohol-associated hepatitis were not included in the multivariable model, as these were co-linear with hospitalized status, and alcohol-associated liver disease, respectively. Our multivariable model for discordant classification utilized case exclusion for any missing variable. The fully-adjusted multivariable model included 478 of 481 patients. Given the extremely limited missing data, we did not perform data imputation in multivariable analysis. All regression analyses were adjusted for center clustering, *using the STATA vce cluster command (*https://www.stata.com/manuals/xtvce_options.pdf*).*

To determine the minimum sample size for the primary concordance analysis using the Cohen's Kappa statistic, we followed best practices as recommended by Rotondi *et al*[10] and Donner *et al*[11] We regarded the minimum acceptable Kappa to be 0.8 (null hypothesis), which would represent substantial agreement. In reference to this, we aimed to have an 80% power to detect a decrease in Kappa of at least 0.1 (alternative hypothesis). We evaluated several a priori proportions of ACLF grade distributions, and determined that across all scenarios a sample size of at least 300 would be sufficient for this purpose. We were also interested in positive predictive value. We calculated that a sample size of 481 patients would provide a confidence interval of ±0.035, targeting a positive predictive value of 0.80.

Statistical analyses were performed using STATA/MP 16.1 (College Station, TX). This study was approved by the Institutional Review Board at each participating center.

RESULTS

Study Population

Among 481 sampled LT registrants across 3 centers from 2013–2019, 250 (52%) had no ACLF, 75 (16%) had ACLF Grade 1, 79 (16%) had ACLF Grade 2, and 77 (16%) had ACLF Grade 3 by UNOS categorization. Among the total cohort, median age was 55 (IQR 48–61), 247 (51%) were non-Hispanic White, and median MELD-Na score was 23 (IQR 16–34). Patient characteristics of the full cohort and by ACLF grade are summarized in Table 1.

By chart review of the full cohort (N=481), 338 (70%) had no ACLF, 28 (6%) had ACLF Grade 1, 39 (8%) had ACLF Grade 2, and 76 (16%) had ACLF Grade 3 by chart review. Among this full cohort, 180 (37%) had an acute liver decompensation, 68 (14%) with acute development of large ascites, 100 (21%) with encephalopathy, 43 (9%) with gastrointestinal bleeding, 110 (23%) with bacterial infection.

Concordance and Accuracy Metrics

The concordance of no ACLF, Grade 1, Grade 2, and Grade 3 by UNOS vs. chart review was 72%, 64%, 56%, and 64% (Table 2). Classification of ACLF grades by UNOS (vs. chart review) had poor positive predictive values (ranging 24–64%), but high negative predictive values (ranging 93–98%). Classification by UNOS (vs. chart review) of liver, coagulation, and kidney failures had high positive and negative predictive values (>90%), but brain, circulation, and circulation failures had poor accuracy metrics. Negative predictive, positive predictive, sensitivity, specificity, and AUC values for ACLF grades, and each organ failure are detailed in Table 3.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

To assess concordance across all ACLF categories, the calculated overall Cohen's kappa coefficient was 0.48 (95%CI 0.42–0.54), suggesting weak agreement.

Factors Associated with Discordant ACLF Classification

Among 481 patients, 148 (31%) had discordant ACLF classification by UNOS vs. chart review. UNOS tended to overestimate ACLF in patients without ACLF and in patients with ACLF grade 1 by chart review: among 338 patients without ACLF by chart review, 94 (28%) were classified as having ACLF by UNOS (Table 2). UNOS tended to underestimate ACLF grade in patients with Grade 2 and 3 ACLF by chart review (Table 2). In multivariable analysis, higher MELD-Na score (aOR 1.06, 95%CI 1.00–1.12) was associated with discordant ACLF classification (Supplemental Table 3). The most common reason for overestimation of ACLF grade by UNOS was absence of an acute decompensation: 92 of 104 (88%) patients with overestimated ACLF by UNOS had no acute development of large ascites, hepatic encephalopathy, gastrointestinal bleeding, or bacterial infection by chart review. The most common reasons for underestimation of ACLF grade by UNOS was due to underestimation of brain failure and circulatory failure: among 44 patients with underestimated ACLF, 1 (2%) had brain failure by UNOS vs. 29 (66%) by chart review, and 0 had circulatory failure by UNOS vs. 18 (41%) by chart review. Liver decompensations and organ failures by UNOS vs. chart review for overestimated and underestimated patients are summarized in Supplemental Table 4.

DISCUSSION

In this retrospective multi-center study, we show that ACLF categorization by UNOS is not concordant with manual chart review. We outline reasons for this discordance, all of which could be modified by UNOS reform by amending existing variables and adding new variables. These findings serve as a cautionary signal in regards to ongoing proposals relying quantitatively on prior UNOS studies that may be providing estimates skewed by misclassification bias. Our study suggests the need for large prospective studies regarding ACLF in LT and UNOS reform to provide valid estimates to more accurately inform changes to allocation policy.

Our study provides important perspective to interpret prior ACLF studies using UNOS as a data source, and how to improve validity for future UNOS studies. Prior UNOS studies[5,7] in ACLF categorize liver, coagulation, and kidney failure by bilirubin, INR, and creatinine/dialysis, respectively, which reflect MELD-Na score. They categorize brain, respiratory, circulatory failure, by coded encephalopathy, mechanical ventilation, and vasopressors, respectively, which reflect hospitalization status. These prior UNOS studies of ACLF have not incorporated hospitalized status into their modeling, and have included both hospitalized and non-hospitalized patients. The co-dependence of ACLF categorization on other clinically important variables can produce unreliable estimates in multivariable models.

We propose two main avenues for UNOS reform to address these deficiencies. First, as UNOS studies do not capture the additional nuances of organ failure categorizations, including oxygen saturation, blood gas levels, reasons for mechanical ventilation,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

gastrointestinal bleeding, bacterial infections, and timing of liver decompensation—these are variables that should be added to UNOS. New acute decompensation is foundational to the ACLF definition, and thus appears to be especially high yield as a new UNOS variable. Second, we found that brain failure and circulatory failure were significantly underestimated by UNOS—data are entered into UNOS by coordinators who rely on documentation in which the presence of encephalopathy or vasopressor support may not be readily obvious, which represents an area that would benefit from increased awareness and education. These results highlight areas for UNOS data reform that would allow the transplant field to advance knowledge of ACLF among LT candidates.

This study had limitations. First, our data are retrospective, and misclassification by chart review is also possible. However, reasons for re-classification were objective with clear validated, and standardized criteria (e.g. new decompensation, SpO2/FiO2 ratio), we had near 100% completeness of data as a reflection of extensive clinical evaluation and documentation surrounding LT candidates at listing, suggesting validity to our data collection methodology. In addition, variables prone to subjectivity (e.g. encephalopathy) are unlikely to affect our results, as organ failures by EASL-CLIF criteria[1] typically represent extreme manifestations (e.g. only grade 3 or 4 encephalopathy constitute brain failure). Second, our sample size is modest and reflects three high volume LT centers, which may limit generalizability. We addressed this limitation by random sampling, adjusting for center clustering, and by providing confidence intervals as a metric for uncertainty. Nevertheless, our findings should be interpreted as preliminary, and emphasize the need for high-quality, larger, prospective studies to investigate the incorporation of ACLF into LT policy. Third, our study does not assess other definitions[12] of ACLF (e.g. APASL, NACSELD) as previous UNOS studies do not focus on those classifications—UNOS does not provide variables that reflect the differences between these definitions, so we would not anticipate UNOS to have improved concordance with either of these alternate ACLF definitions. Fourth, our study cohort was purposely enriched with ACLF and does not provide information regarding prevalence of ACLF in a transplant population. Finally, our study did not assess treatments for acute liver decompensation, longitudinal clinical factors on the waitlist, or post-LT outcomes, which are areas for future investigation to advance knowledge of ACLF in LT.

In conclusion, we show that ACLF categorization by UNOS is not concordant with manual chart review and provide insights to rectify this discordance. These findings highlight the value in encouraging prospective studies regarding ACLF in LT and UNOS reform.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

Jennifer Dodge, MPH, for her statistical input.

Funding Sources: None

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Abbreviations:

ACLF	acute-on-chronic-liver-failure
AH	alcohol-associated hepatitis
AUC	Area Under Curve
LT	liver transplantation
NPV	negative predictive value
PPV	positive predictive value
UNOS	United Network for Organ Sharing

References

1. Moreau R, Jalan R, Gines P, Pavesi M, Angeli P, Cordoba J, et al. Acute-on-chronic liver failure is a distinct syndrome that develops in patients with acute decompensation of cirrhosis. Gastroenterology 2013;144:1426–37, 37 e1–9. [PubMed: 23474284]
2. Hernaez R, Liu Y, Kramer JR, Rana A, El-Serag HB, Kanwal F. Model for end-stage liver disease-sodium underestimates 90-day mortality risk in patients with acute-on-chronic liver failure. J Hepatol 2020;73:1425–33. [PubMed: 32531416]
3. Sundaram V, Jalan R. Waiting list priority for patients with acute-on-chronic liver failure: Not just horseplay. Liver Transpl 2021. E-pub ahead of print.
4. Belli LS, Duvoux C, Artzner T, Bernal W, Conti S, Cortesi PA, et al. Liver transplantation for patients with acute-on-chronic liver failure (ACLF) in Europe: Results of the ELITA/EF-CLIF collaborative study (ECLIS). J Hepatol 2021;75:610–22. [PubMed: 33951535]
5. Sundaram V, Shah P, Wong RJ, Karvellas CJ, Fortune BE, Mahmud N, et al. Patients With Acute on Chronic Liver Failure Grade 3 Have Greater 14-Day Waitlist Mortality Than Status-1a Patients. Hepatology 2019;70:334–45. [PubMed: 30908660]
6. Sundaram V, Jalan R, Ahn JC, Charlton MR, Goldberg DS, Karvellas CJ, et al. Class III obesity is a risk factor for the development of acute-on-chronic liver failure in patients with decompensated cirrhosis. J Hepatol 2018;69:617–25. [PubMed: 29709681]
7. 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;156:1381–91 e3.
8. Crabb DW, Bataller R, Chalasani NP, Kamath PS, Lucey M, Mathurin P, et al. Standard Definitions and Common Data Elements for Clinical Trials in Patients With Alcoholic Hepatitis: Recommendation From the NIAAA Alcoholic Hepatitis Consortia. Gastroenterology 2016;150:785–90. [PubMed: 26921783]
9. Mchugh M. Interrater reliability: the kappa statistic. Biochemia Medica 2012;22:276–82. [PubMed: 23092060]
10. Rotondi M, Donner A. A Confidence Interval Approach to Sample Size Estimation for Interobserver Agreement Studies with Multiple Raters and Outcomes. Journal of Clinical Epidemiology 2012;65:778–84. [PubMed: 22560852]
11. Donner A. Sample Size Requirements for Interval Estimation of the Intraclass Kappa Statistic. Communication in Statistics 1999;28:415–29.
12. Zaccherini G, Weiss E, Moreau R. Acute-on-chronic liver failure: Definitions, pathophysiology and principles of treatment. JHEP Rep 2021;3:100176.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

- UNOS studies on ACLF are informing global transplant allocation policies, but the validity of these data is unknown.
- We found that ACLF identified by UNOS is only weakly concordant with chart review.
- Classification of acute decompensation and brain/respiratory failures are common sources of discordance.
- These data highlight the importance of prospective studies regarding ACLF in LT and should encourage UNOS reform.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1.

Patient Characteristics by ACLF Grade

		ACLF Classification by UNOS			
	Total (N=481)	**None (N=250)**	**Grade 1 (N=75)**	**Grade 2 (N=79)**	**Grade 3 (N=77)**
Characteristics					
Male, n (%)	280 (58)	147 (59)	44 (59)	47 (60)	42 (54)
Age, median (IQR)	55 (48–61)	56 (48–61)	57 (49–62)	55 (47–61)	53 (46–60)
Non-Hispanic White, n (%)	247 (51)	143 (57)	39 (52)	32 (40)	33 (43)
Primary Listing Diagnosis, n (%)					
Alcohol	155 (32)	71 (28)	27 (36)	27 (34)	30 (39)
NASH/Cryptogenic	103 (21)	46 (18)	21 (28)	22 (28)	14 (18)
HBV/HCV	118 (24)	65 (26)	16 (21)	19 (24)	18 (23)
AIH/PBC/PSC	69 (14)	44 (18)	7 (9)	7 (9)	11 (14)
Other	36 (7)	24 (10)	4 (5)	4 (5)	4 (5)
Alcohol-Associated Hepatitis by NIAAA Criteria, n (%)	19 (4)	7 (3)	5 (7)	3 (4)	4 (5)
Primary Insurance, n (%)					
Private	163 (34)	81 (32)	31 (41)	30 (38)	21 (27)
Medicare	157 (33)	89 (36)	24 (32)	23 (29)	21 (27)
Medicaid	161 (33)	80 (32)	20 (27)	26 (33)	35 (45)
Hospitalized, n (%)	218 (45)	39 (16)	42 (56)	64 (81)	73 (95)
In Intensive Care Unit, n (%)	135 (28)	9 (4)	16 (21)	41 (53)	69 (90)
Diabetes, n (%)	131 (27)	53 (21)	36 (48)	22 (28)	20 (26)
MELD-Na score, median (IQR)	23 (16–34)	16 (14–20)	28 (22–33)	35 (31–39)	43 (40–45)
Organ Failures					
Liver	145 (30)	9 (4)	19 (25)	48 (61)	69 (90)
Coagulation	127 (26)	2 (1)	10 (13)	45 (57)	70 (91)
Kidney	167 (35)	0 (0)	44 (59)	49 (62)	74 (96)
Brain	51 (11)	8 (3)	2 (3)	16 (20)	25 (32)
Respiratory	19 (4)	0 (0)	0 (0)	0 (0)	19 (25)
Circulatory	21 (4)	0 (0)	0 (0)	0 (0)	21 (27)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

Concordance of ACLF Grade by UNOS vs. chart review

ACLF Grade		By UNOS				
		0	**1**	**2**	**3**	**Total**
By Chart Review	**0**	**244 (72)**	41 (12)	28 (8)	25 (7)	338 (100)
	1	2 (7)	**18 (64)**	7 (25)	1 (4)	28 (100)
	2	2 (5)	13 (33)	**22 (56)**	2 (5)	39 (100)
	3	2 (3)	3 (4)	22 (29)	**49 (64)**	76 (100)
	Total	250 (52)	75 (16)	79 (16)	77 (16)	481 (100)

Underestimation by UNOS
Overestimation by UNOS

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 3.

Accuracy Metrics for ACLF Grade and Organ Failure by UNOS vs. Chart Review

		Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	AUC (95% CI)
ACLF	**None**	0.96 (0.91–0.98)	0.72 (0.67–0.77)	0.59 (0.53–0.66)	0.98 (0.95–0.99)	0.84 (0.81–0.87)
	Grade 1	0.64 (0.44–0.81)	0.87 (0.84–0.90)	0.24 (0.15–0.35)	0.98 (0.96–0.99)	0.76 (0.67–0.85)
	Grade 2	0.56 (0.40–0.72)	0.87 (0.84–0.90)	0.28 (0.18–0.39)	0.96 (0.93–0.98)	0.72 (0.64–0.80)
	Grade 3	0.64 (0.53–0.75)	0.93 (0.90–0.95)	0.64 (0.52–0.74)	0.93 (0.90–0.96)	0.79 (0.73–0.84)
Organ Failure	**Liver**	0.96 (0.92–0.99)	0.98 (0.95–0.99)	0.94 (0.89–0.98)	0.99 (0.97–1.00)	0.97 (0.95–0.99)
	Coagulation	0.93 (0.88–0.97)	0.96 (0.94–0.98)	0.90 (0.83–0.94)	0.98 (0.96–0.99)	0.95 (0.93–0.97)
	Kidney	0.94 (0.89–0.97)	0.98 (0.96–1.00)	0.97 (0.93–0.99)	0.96 (0.94–0.98)	0.96 (0.94–0.98)
	Brain	0.13 (0.09–0.18)	0.92 (0.88–0.95)	0.63 (0.48–0.76)	0.51 (0.46–0.56)	0.53 (0.50–0.55)
	Respiratory	0.36 (0.21–0.53)	0.99 (0.97–1.00)	0.74 (0.49–0.91)	0.95 (0.92–0.96)	0.67 (0.60–0.75)
	Circulatory	0.15 (0.08–0.25)	0.98 (0.96–0.99)	0.52 (0.30–0.74)	0.86 (0.83–0.89)	0.56 (0.52–0.60)

NPV: negative predictive value; PPV: positive predictive value; AUC: area under curve