



Published in final edited form as:

*Clin Gastroenterol Hepatol*. 2022 May ; 20(5): e1121–e1134. doi:10.1016/j.cgh.2021.06.050.

## Risk Prediction Models for Post-Operative Decompensation and Infection in Patients with Cirrhosis: A Veterans Affairs Cohort Study

Nadim Mahmud, MD MS MPH MSCE<sup>1,2,3,4</sup>, Zachary Fricker, MD<sup>5</sup>, James D. Lewis, MD MSCE<sup>1,4</sup>, Tamar H. Taddei, MD<sup>6,7</sup>, David S. Goldberg, MD MSCE<sup>8</sup>, David E. Kaplan, MD MSc<sup>1,3</sup>

1. Division of Gastroenterology and Hepatology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA
2. Leonard David Institute of Health Economics, University of Pennsylvania, Philadelphia, PA
3. Department of Medicine, Corporal Michael J. Crescenz VA Medical Center, Philadelphia, PA
4. Center for Clinical Epidemiology and Biostatistics, Department of Biostatistics, Epidemiology & Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA
5. Division of Gastroenterology, Hepatology, and Nutrition, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA
6. Division of Digestive Diseases, Yale University School of Medicine, New Haven, CT
7. VA Connecticut Healthcare System, West Haven, CT
8. Division of Digestive Health and Liver Diseases, Department of Medicine, University of Miami Miller School of Medicine, Miami, FL

### Abstract

**Background and Aims:** Patients with cirrhosis have an increased risk of post-operative mortality for a range of surgeries, however they are also at risk of post-operative complications such as infection and cirrhosis decompensation. To date there are no prediction scores that specifically risk stratify patients for these morbidities.

**Corresponding Author:** Nadim Mahmud MD MS MPH MSCE, 3400 Civic Center Boulevard, 4<sup>th</sup> Floor, South Pavilion, Philadelphia, PA 19104, Phone: 215-349-8222, Fax: 215-349-5915, nadim@pennmedicine.upenn.edu.

Author Contributions:

Intellectual genesis: NM, ZF, JDL, THT, DSG, DEK

Data preparation and management: NM, ZF, DEK

Formal statistical analysis: NM

Visualization: NM

Manuscript writing: NM

Critical manuscript review and editing: NM, ZF, JDL, THT, DSG, DEK

**Disclosures:** The authors have no additional disclosures or conflicts as relevant to this manuscript.

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

**Methods:** This was a retrospective study using data of patients with cirrhosis who underwent diverse surgeries in the Veterans Health Administration. Validated algorithms and/or manual adjudication were used to ascertain post-operative decompensation and post-operative infection through 90 days. Multivariable logistic regression was used to evaluate prediction models in derivation and validation sets using variables from the recently-published VOCAL-Penn cirrhosis surgical risk scores for post-operative mortality. Models were compared to the Mayo risk score, MELD-sodium, and Child-Turcotte-Pugh (CTP) scores.

**Results:** A total 4,712 surgeries were included; patients were predominantly male (97.2%), white (63.3%), and with alcohol-related liver disease (35.3%). Through 90 post-operative days, 8.7% of patients experienced interval decompensation, and 4.5% infection. Novel VOCAL-Penn prediction models for decompensation demonstrated good discrimination for interval decompensation (C-statistic 0.762 vs. 0.663 Mayo vs. 0.603 MELD-sodium vs. 0.560 CTP,  $p < 0.001$ ), however discrimination was only fair for post-operative infection (C-statistic 0.666 vs. 0.592 Mayo [ $p = 0.13$ ] vs. 0.502 MELD-sodium [ $p < 0.001$ ] vs. 0.503 CTP [ $p < 0.001$ ]). The model for interval decompensation had excellent calibration in both derivation and validation sets.

**Conclusion:** We report the derivation and internal validation of a novel, parsimonious prediction model for post-operative decompensation in patients with cirrhosis. This score demonstrated superior discrimination and calibration as compared to existing clinical standards, and will be available at [www.vocalpennscore.com](http://www.vocalpennscore.com).

## Keywords

cirrhosis surgical risk; decompensation; infection; VOCAL-Penn score; Veterans Affairs

## Introduction

The burden of cirrhosis in the United States is on the rise, owing to increases in non-alcoholic fatty liver disease as well as alcohol use disorder.<sup>1, 2</sup> In tandem with these epidemiologic shifts, a growing volume of patients with cirrhosis are undergoing major surgeries on an annual basis.<sup>3</sup> Given that cirrhosis confers an increased risk of post-operative mortality,<sup>4</sup> it is critical to accurately risk stratify patients for such procedures to inform patient and surgical decision making. Numerous prior prediction models have been designed to estimate the risk of post-operative mortality. This includes the model for end-stage liver disease-sodium (MELD-Na) score, Child-Turcotte-Pugh (CTP) score, the Mayo risk score,<sup>5</sup> and the recently developed VOCAL-Penn cirrhosis surgical risk score ([www.vocalpennscore.com](http://www.vocalpennscore.com)).<sup>6, 7</sup> Although mortality is clearly an important patient outcome, the development of post-operative morbidity such as interval cirrhosis decompensation and post-operative infections are outcomes that also might impact the balance of risk and benefit of proceeding with an individual surgery. These events mediate an increased risk of post-operative death,<sup>8, 9</sup> but more immediately, they impact patient functioning, quality of life, and healthcare resource utilization.<sup>10</sup> Thus, reliable prognostic models for these outcomes would be of immediate pragmatic clinical value to both patients and providers.

In this study, our goal was to build from the recently-published VOCAL-Penn cirrhosis surgical risk score and extend prediction models to two post-operative morbidity outcomes through 90 days: interval cirrhosis decompensation and post-operative infection.

## Methods

### Study Design and Data Source

This was a retrospective cohort study using a merged dataset from the Veterans Outcomes and Costs Associated with Liver Disease (VOCAL) and Veterans Affairs Surgical Quality Improvement Program (VASQIP) cohorts. The derivation and merging of these sources has been described previously.<sup>6</sup> In brief, VOCAL contains granular longitudinal data from over 129,000 patients with cirrhosis identified in the Veterans Health Administration (VHA) between January 1, 2008 and December 31, 2015.<sup>11, 12</sup> VASQIP is a manually-adjudicated surgical dataset in the VHA initially developed to support peri-operative quality improvement research.<sup>13</sup> The merged VOCAL-VASQIP dataset, with data on 4,712 unique surgeries, was previously used to derive and internally validate the VOCAL-Penn cirrhosis surgical risk score.<sup>6</sup> In this study, the same data were used for the purposes of evaluating VOCAL-Penn predictors for post-operative decompensation or infection. This included patients age >18 years with established cirrhosis who underwent a surgery of interest (defined below). Patients who underwent prior liver transplantation were excluded.

### Surgery Categories

Consistent with prior methods in the derivation and external validation of the VOCAL-Penn cirrhosis surgical risk score,<sup>6, 7</sup> we focused on six surgery categories: abdominal wall (e.g., umbilical or inguinal hernia repair), vascular (e.g., femoral-femoral bypass graft), major orthopedic (e.g., total hip arthroplasty), chest/cardiac (e.g., pulmonary lobectomy, coronary artery bypass graft), open abdominal, and laparoscopic abdominal (e.g., cholecystectomy, colectomy, appendectomy). These categories were established based on Current Procedure Terminology (CPT) codes as defined in Supplemental Table 1, where specific surgery examples are also provided.

### Variable Collection

For each surgery, we collected patient demographics (age, sex, race), smoking history (never, former, current), body mass index (BMI), and comorbidities (hypertension, diabetes, obesity [BMI ≥ 30], atrial fibrillation, coronary artery disease, congestive heart failure). Prior cirrhosis decompensations were determined using a previously validated algorithm in the VHA system,<sup>14</sup> and pre-operative transjugular intrahepatic portosystemic shunt (TIPS) placement was determined using CPT codes (37182, 37183). Etiology of liver disease was classified as hepatitis C virus (HCV), hepatitis B virus, alcohol-related liver disease (ALD), HCV + ALD, non-alcoholic fatty liver disease (NAFLD), or other (including autoimmune hepatitis, primary biliary cholangitis, primary sclerosing cholangitis), consistent with a previously-validated approach.<sup>15</sup> The following pre-operative laboratory data within 30 days of surgery were obtained: aspartate aminotransferase (AST), alanine aminotransferase (ALT), total bilirubin, albumin, platelet count, and international normalized ratio (INR). Finally, adjudicated VASQIP data were used to obtain American

Society of Anesthesiologists (ASA) physical status classification (2, 3, 4) as well as emergency status of the surgery. Importantly, the decision to categorize ASA as 2, 3, or 4 as opposed to 3 for compensated and 4 for decompensated cirrhosis (as was done in the Mayo risk score) has been evaluated in detail in prior work from our group.<sup>6</sup> The above data were used to compute prediction scores to be evaluated in this study (Supplemental Table 2), including the Mayo risk, MELD-Na, and CTP (A, B, C) scores. For subjective components of the CTP score, ascites and hepatic encephalopathy were determined using a previously validated VHA algorithm.<sup>11</sup>

## Outcomes Ascertainment

The primary outcomes of interest were (1) interval cirrhosis decompensation within 90 days of surgery and (2) infection within 30 days of surgery (as adjudicated post-operative infection data are only available through 30 days in VASQIP). Regarding the first outcome, we used a previously-validated VHA algorithm to identify the combined endpoint of portal-hypertension related decompensation, clinically-significant ascites, hepatic encephalopathy, hepatorenal syndrome, or variceal bleeding.<sup>11, 14</sup> These algorithms rely on combinations of CPT codes, International Classification of Diseases (ICD) codes, and pharmacy/medication administration data, and have been extensively utilized in the VHA cirrhosis literature.<sup>12, 16–19</sup> Only patients with clinically-significant manifestations of portal hypertension were ascertained in this manner (i.e. mild ascites was not considered a decompensation, but moderate/severe ascites was).<sup>20</sup> Regarding the second outcome, in the VASQIP database selected post-operative complications are manually adjudicated by trained nurse data managers through 30 days after surgery. We used these data to create a combined endpoint for post-operative infection, which included urinary tract infection, pneumonia, wound superinfection, skin/soft tissue infection, or sepsis. Of note, because the algorithm to determine clinically significant ascites includes administrative codes for spontaneous bacterial peritonitis (SBP), SBP events could be classified as both interval cirrhosis decompensation and infection in this study.

## Statistical Analysis

Descriptive data for the cohort were computed as medians and interquartile ranges (IQRs, displayed as 25<sup>th</sup> and 75<sup>th</sup> percentiles in results) for continuous data and as percentages for categorical data. Univariable associations between surgery category and the primary outcomes and between the primary outcomes and 90-day mortality were tested using the chi square test. For the purposes of subsequent prediction modeling, the dataset was divided into 80% derivation and 20% validation subsets using a random number generator. Logistic regression was used to derive prediction models for each outcome. *A priori* variables considered in models included the predictors used in the VOCAL-Penn cirrhosis surgical risk score: age, albumin, total bilirubin, platelet count, obesity (body mass index ≥ 30), non-alcoholic fatty liver disease, ASA score, and surgery category. The rationale for focusing on these variables was twofold. First, these are established predictors of post-operative mortality in patients with cirrhosis,<sup>5, 6, 21, 22</sup> and post-operative decompensation and infection are likely mediators of mortality.<sup>8, 9</sup> Second, given the goal of extending the applications of the VOCAL-Penn tool, it was desirable to utilize a parsimonious set of variables to maximize clinical adoption. Prior to multivariable analyses, univariable

analyses in the derivation set were conducted. Continuous predictors were evaluated for non-linearity with each outcome using locally-weighted scatterplot smoothing curves. Based on this analysis, restricted cubic splines were used to transform platelet count, albumin, and total bilirubin. Knot specifications are given in Supplemental Table 3. In subsequent multivariable analysis, we began with a full model and tested multiple reduced models using a purposeful selection approach. Different candidate models were evaluated on the basis of minimized Bayesian Information Criterion (BIC), and maximized model discrimination. Full models for both outcomes were ultimately selected based on these considerations. Finally, we considered the possible impact of pre-operative TIPS on each outcome by adding this covariate to the final model and comparing models using the likelihood ratio test. An alpha threshold of 5% was used to determine statistical significance.

The predictive performance of models was evaluated in terms of discrimination and calibration, in both derivation and validation sets. Receiver operating characteristic (ROC) curves with computation of C-statistics and 95% confidence intervals (CIs) were used to evaluate discrimination. Smoothed plots of observed events versus predicted probabilities were used to evaluate calibration by visual inspection, consistent with guideline recommendations.<sup>23</sup> Perfect calibration would be indicated by points lying along a 45-degree line. The above final models (called “VOCAL-Penn Refit”) were compared to several established risk prediction scores in liver disease: the VOCAL-Penn cirrhosis surgical risk score for post-operative mortality (called “VOCAL-Penn Mortality” in tables/figures, 90-day mortality prediction for decompensation models, 30-day mortality prediction for infection models), the Mayo surgical risk score (90-day mortality prediction for decompensation models, 30-day mortality prediction for infection models), MELD-Na, and the CTP score. Model discrimination was compared globally using a 5% alpha threshold, and in pairwise comparisons using the Bonferroni adjustment (1.25% alpha threshold for significance).

## Sensitivity Analysis

Given the possibility that prediction model performance for post-operative decompensation might vary on the basis of prior cirrhosis decompensation or risk category of a surgery, we performed two sensitivity analyses. First, we evaluated model discrimination among (1) patients with previously-compensated cirrhosis and (2) patients with a history of prior decompensation. Second, we dichotomized surgery type into the three highest mortality (open abdominal, chest/cardiac, abdominal wall) and three lowest mortality groups (laparoscopic abdominal, vascular, major orthopedic), and evaluated model discrimination separately for each group.

## Results

### Cohort Characteristics

The cohort of patients undergoing 4,712 unique surgeries had a median age 64 years, were 97.2% male, and 63.3% white (Table 1). The most common surgery category was abdominal wall (27.8%), followed by major orthopedic (27.5%). The most common etiologies of liver disease were ALD (35.3%) and HCV + ALD (29.5%). The majority of patients were CTP A (88.3%) and ASA class 3 (67.8%). Through 90 post-operative days, 408 (8.7%)

patients experienced interval cirrhosis decompensation, and through 30 post-operative days 244 (5.2%) developed an infection. Median time to post-operative decompensation was 11 days (IQR 4, 29), and median time to post-operative infection was 10 days (IQR 5, 16; Supplemental Table 4). Each outcome was significantly associated with surgery category (each  $p < 0.001$ , Table 2), with open abdominal surgery conferring the highest unadjusted risk overall (17.9% decompensation and 10.5% infection). By contrast, vascular surgeries had the lowest risk of decompensation (4.7%) and abdominal wall surgeries had the lowest risk of infection (2.8%). Of the 408 patients with decompensation events, 199 (48.8%) patients developed interval ascites, 121 (29.7%) hepatic encephalopathy, 84 (20.6%) hepatorenal syndrome, and 72 (17.6%) variceal bleeding.

### Prediction Models for Post-operative Decompensation

In multivariable analysis, increasing ASA class was associated with a higher odds of post-operative decompensation (OR 1.36, 95% CI 1.05 – 1.76,  $p = 0.02$ ), and open abdominal surgeries had the highest odds relative to laparoscopic abdominal surgeries (OR 1.84, 95% CI 1.12 – 3.02; Table 3). Lower platelet count and albumin were associated with increased odds of post-operative decompensation, as was higher total bilirubin. The addition of pre-operative TIPS as a covariate did not improve model fit ( $p = 0.61$ ). In terms of prediction model performance, the discrimination of the final model (VOCAL-Penn Refit) was very good in both derivation and validation sets (C-statistics 0.766 and 0.762, respectively), and significantly higher as compared to VOCAL-Penn Mortality, Mayo score, MELD-Na, and CTP score models (each  $p < 0.0125$ ; Figure 1A/B, Table 4). The calibration of the VOCAL-Penn Refit model was excellent across the spectrum of risk in both derivation and validation sets (Figure 1C). By contrast, the Mayo risk score significantly overestimated post-operative decompensation across the risk spectrum.

In sensitivity analyses, we found that model discrimination for 90-day post-operative decompensation for VOCAL-Penn Refit remained good when stratified by prior compensation/decompensation status, though was higher among patients with previously compensated cirrhosis versus previously decompensated cirrhosis (C-statistic 0.798 vs. 0.705; Supplemental Table 5). The VOCAL-Penn Refit model demonstrated numerically higher discrimination than all other scores. Similar findings were observed when analyses were limited to higher-mortality and lower-mortality surgery categories (Supplemental Table 6).

### Prediction Models for Post-operative Infection

In the final multivariable model for post-operative infection, surgery category and albumin were significantly associated with the outcome (each joint  $p < 0.001$ ; Table 3). The addition of the pre-operative TIPS covariate again did not improve model fit ( $p = 0.83$ ). The discrimination of the final model (VOCAL-Penn Refit) for 30-day post-operative infection was fair in both derivation and validation set (C-statistic 0.692 and 0.654, respectively; Figure 2A/B, Table 4). Although this was significantly higher than all competing models in the derivation set (each  $p < 0.0125$ ), the VOCAL-Penn Refit model was similar to VOCAL-Penn Mortality, Mayo risk, and MELD-Na scores in the validation set ( $p = 0.49$ ,  $p = 0.30$ , and  $p = 0.046$ , respectively). The calibration of the VOCAL-Penn Refit model was good in



the derivation cohort by visual inspection, though there was some evidence of regional overestimation and underestimation of risk in the validation cohort (Figure 2C).

## Discussion

In this national study of 4,712 diverse surgeries performed on patients with cirrhosis, we identified prediction models for post-operative morbidity outcomes through 90 days. When refit, the covariates in the VOCAL-Penn cirrhosis surgical risk scores demonstrated good discrimination and excellent calibration in predicting post-operative hepatic decompensation. An analogous model demonstrated only fair prediction for post-operative infection, however it is noteworthy that both models represent significant improvements in prognostication as compared to established risk scores in the field of chronic liver disease.

The most impactful finding in this study is improved prediction of post-operative cirrhosis decompensation events. Prior studies have identified variables associated with post-operative decompensation, however these studies are limited in terms of sample size and/or focus on individual surgery types. For example, Nyberg et al. studied 853 patients with cirrhosis who underwent orthopedic surgeries and identified lower serum albumin as being associated with 90-day decompensation.<sup>24</sup> Suman et al. conducted a study of 44 patients with cirrhosis undergoing major cardiac surgery and found that low albumin and elevated bilirubin were associated with decompensation.<sup>9</sup> Higher MELD and CTP scores were also associated with this outcome. Our findings are consistent with these studies, however they expand on existing literature in several important ways. First, prior studies were focused on etiology-based models, whereas the refitted VOCAL-Penn model is designed for prediction. Second, the VHA cohort contains a much larger sample size and addresses multiple broad surgery categories. This will serve to expand the clinical utility of the prediction score. Third, the VOCAL-Penn model was compared to existing clinical standards such as the MELD-Na, CTP, and Mayo risk score, and it demonstrated substantially improved predictive value. Fourth, an additional benefit of using the VOCAL-Penn model is that discrete probabilities of post-operative decompensation are provided, in contrast to risk scores such as MELD-Na and CTP. Fifth, because of the focus on modeling using VOCAL-Penn Mortality covariates, the refitted VOCAL-Penn score for post-operative decompensation remains parsimonious and can be computed with the same inputs (to be available in conjunction with post-operative mortality predictions at [www.vocalpennscore.com](http://www.vocalpennscore.com)). Finally, it is important to highlight the clinical impact of the refitted VOCAL-Penn score for post-operative decompensation. In addition to affording more prognostic depth in conversations between patients and providers, the score may improve post-operative triage decisions and allocation of healthcare resources. For example, patients predicted to have a high risk of post-operative decompensation may warrant closer post-operative monitoring, such as in an intensive care or stepdown ward, whereas low-risk patients may be observed in the general medical ward or even in the outpatient setting. Accurate risk stratification may also facilitate patient selection for research studies testing interventions to mitigate post-operative complications in high-risk individuals.

Although the VOCAL-Penn prediction model for post-operative infection demonstrated superior discrimination as compared to other scores in the derivation set, there was no difference as compared to the Mayo or MELD-Na score in the validation set. Furthermore, the overall discrimination was only fair with a C-statistic of only 0.654. Based on these findings, we would not advocate for the use of any of the prediction models evaluated in this study for prediction of post-operative infection. It is interesting that we found a discordance in predictive performance between VOCAL-Penn models and the outcomes of decompensation and infection. Given the strength of association between model covariates and post-operative decompensation and mortality, this would suggest that cirrhosis decompensation events after surgery are a stronger mediator of mortality than post-operative infection. It is also likely that post-operative infection may be predicted by distinct variables that are not captured in the VOCAL-Penn models, such as hospital volume, surgeon expertise, rural versus urban setting, or individual markers of relative immunocompromise. Future studies are needed to better identify predictors of post-operative infection and reversible risk factors for this complication.

There are important limitations to this study that we acknowledge. First, there is possible misclassification of exposures and outcomes. To minimize this issue, we used widely-validated algorithms for exposure variables wherever possible. Regarding outcomes misclassification, we expect this to be minimal though we acknowledge that selected decompensations such as ascites and hepatic encephalopathy involve subjectivity. Additionally, the VASQIP dataset is not inclusive of all types of infection, and thus the estimates reported in this study may be conservative. Second, there are clear external validity limitations to this study given the nature of the VHA cohort. In particular, the patients are predominantly male and enriched in psychosocial comorbidities relative to the general population. External validation of the refitted VOCAL-Penn model would broaden its usability to healthcare settings outside the VHA. Third, as the cohort only contained patients who underwent surgery, we are not able to predict potential outcomes for patients who were deemed to be poor surgical candidates on prior clinical grounds. Indeed, patients in this study were generally low MELD, CTP A, and ASA 3, and ultimately patients were likely selected for surgery using heterogeneous clinical approaches. Thus, the VOCAL-Penn scores should only be applied as an adjunct to clinical judgment and not as a replacement. That is, if a patient is clearly not fit for surgery on clinical grounds, the VOCAL-Penn scores should not be applied. If risk aversion surrounding cirrhosis surgeries is relaxed, the models may need to be periodically recalibrated. Fourth, there are surgery categories that are not represented in this study, most notably hepatic resections such as those performed for hepatocellular carcinoma. Future studies evaluating prediction score performance in this specific context would be of interest, and is an area of ongoing research in our group. Finally, we did not find TIPS to be associated with the outcomes in adjusted models. However, given the small number of patients in this cohort who had pre-operative TIPS (only 1.3%) and we could not ascertain the indication for TIPS placement, this study was likely underpowered to meaningfully study this exposure in detail. Thus it remains unclear if pre-operative TIPS might decrease post-operative complications. Similarly, there are additional pre-operative variables which may be salient predictors of post-operative outcomes, including treatment status of viral hepatitis, adequate control of autoimmune



hepatitis, active alcohol use, prophylactic antibiotics, and use of medications such as non-selective beta blockers and statins, which were unfortunately not readily available for study in this cohort. These exposures may be evaluated in detail in future studies.

In conclusion, post-operative decompensation and infection are common events in patients with cirrhosis. We report the derivation and internal validation of a prediction score for 90-day post-operative decompensation that adds significant value beyond existing risk scores applied in the peri-operative setting. This risk score, which is an extension of the VOCAL-Penn cirrhosis surgical risk tool, may be used to augment clinical judgment regarding the decision to proceed with surgery and to clarify prognosis and expected post-operative course in discussion with patients.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

Funding Source:

Nadim Mahmud is supported by an American College of Gastroenterology Junior Faculty Development Award (ACG-JR-010-2020).

David Goldberg has received support from Gilead, Merck, and AbbVie unrelated to the topic of this manuscript. He is also supported by a National Institutes of Health R01 (DK120561).

David E. Kaplan has received support from Gilead, Glycotest and Bayer unrelated to the topic of this manuscript. He is also supported by VA Merit Grants (I01-CX-001933, I01-CX-002010).

Zachary Fricker is supported by the AASLD Foundation Advanced/Transplant Hepatology Award.

Tamar H. Taddei has received support from Bayer unrelated to the topic of this manuscript. She is also supported by a VA Merit Grant (I01-CX-002010).

## Abbreviations:

<b>ALD</b>	alcohol-related liver disease
<b>ALT</b>	alanine aminotransferase
<b>ASA</b>	American Society of Anesthesiologists
<b>AST</b>	aspartate aminotransferase
<b>BIC</b>	Bayesian Information Criterion
<b>BMI</b>	body mass index
<b>CTP</b>	Current Procedure Terminology
<b>CTP</b>	Child-Turcotte-Pugh
<b>HCV</b>	hepatitis C virus
<b>IQR</b>	interquartile range

<b>MELD-Na</b>	Model for End-stage Liver Disease-Sodium
<b>NAFLD</b>	non-alcoholic fatty liver disease
<b>ROC</b>	receiver operating characteristic
<b>TIPS</b>	transjugular intrahepatic portosystemic shunt
<b>VASQIP</b>	Veterans Affairs Surgical Quality Improvement Program
<b>VHA</b>	Veterans Health Administration
<b>VOCAL</b>	Veterans Outcomes and Costs Associated with Liver Disease

## References

1. Kabbany MN, Selvakumar PKC, Watt K, et al. Prevalence of nonalcoholic steatohepatitis-associated cirrhosis in the United States: an analysis of national health and nutrition examination survey data. *American Journal of Gastroenterology* 2017;112:581–587.
2. Mellinger JL, Shedden K, Winder GS, et al. The high burden of alcoholic cirrhosis in privately insured persons in the United States. *Hepatology* 2018;68:872–882. [PubMed: 29579356]
3. Tessitore KM, Mahmud N. Trends in surgical volume and in-hospital mortality among United States cirrhosis hospitalizations. *Annals of gastroenterology* 2021;34:85. [PubMed: 33414627]
4. Friedman LS. The risk of surgery in patients with liver disease. *Hepatology* 1999;29:1617–1623. [PubMed: 10347099]
5. Teh SH, Nagorney DM, Stevens SR, et al. Risk factors for mortality after surgery in patients with cirrhosis. *Gastroenterology* 2007;132:1261–1269. [PubMed: 17408652]
6. Mahmud N, Fricker Z, Hubbard RA, et al. Risk Prediction Models for Post-Operative Mortality in Patients With Cirrhosis. *Hepatology* 2020.
7. Mahmud N, Fricker Z, Panchal S, et al. External Validation of the VOCAL-Penn Cirrhosis Surgical Risk Score in Two Large, Independent Health Systems. *Liver Transplantation* 2021.
8. del Olmo JA, Flor-Lorente B, Flor-Civera B, et al. Risk factors for nonhepatic surgery in patients with cirrhosis. *World journal of surgery* 2003;27:647–652. [PubMed: 12732995]
9. Suman A, Barnes DS, Zein NN, et al. Predicting outcome after cardiac surgery in patients with cirrhosis: a comparison of Child-Pugh and MELD scores. *Clinical Gastroenterology and Hepatology* 2004;2:719–723. [PubMed: 15290666]
10. Scaglione SJ, Metcalfe L, Kliethermes S, et al. Early hospital readmissions and mortality in patients with decompensated cirrhosis enrolled in a large national health insurance administrative database. *Journal of clinical gastroenterology* 2017;51:839–844. [PubMed: 28383303]
11. Kaplan DE, Dai F, Aytaman A, et al. Development and performance of an algorithm to estimate the Child-Turcotte-Pugh score from a national electronic healthcare database. *Clinical Gastroenterology and Hepatology* 2015;13:2333–2341. e6. [PubMed: 26188137]
12. Mahmud N, Kaplan DE, Taddei TH, et al. Incidence and mortality of acute-on-chronic liver failure using two definitions in patients with compensated cirrhosis. *Hepatology* 2019;69:2150–2163. [PubMed: 30615211]
13. Massarweh NN, Kaji AH, Itani KM. Practical guide to surgical data sets: veterans affairs surgical quality improvement program (VASQIP). *JAMA surgery* 2018;153:768–769. [PubMed: 29617543]
14. Re III VL, Lim JK, Goetz MB, et al. Validity of diagnostic codes and liver-related laboratory abnormalities to identify hepatic decompensation events in the Veterans Aging Cohort Study. *Pharmacoepidemiology and drug safety* 2011;20:689–699. [PubMed: 21626605]
15. 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;149:1471–1482. e5. [PubMed: 26255044]

16. Mahmud N, Sundaram V, Kaplan DE, et al. Grade 1 acute on chronic liver failure is a predictor for subsequent grade 3 failure. *Hepatology* 2020;72:230–239. [PubMed: 31677284]
17. Mahmud N, Hubbard RA, Kaplan DE, et al. Declining cirrhosis hospitalizations in the wake of the COVID-19 pandemic: a national cohort study. *Gastroenterology* 2020;159:1134–1136. e3. [PubMed: 32387493]
18. Kanneganti M, Mahmud N, Kaplan DE, et al. Survival benefit of liver transplantation for hepatocellular carcinoma. *Transplantation* 2020;104:104–112. [PubMed: 31283688]
19. Serper M, Weinberg EM, Cohen JB, et al. Mortality and hepatic decompensation in patients with cirrhosis and atrial fibrillation treated with anticoagulation. *Hepatology* 2021;73:219–232. [PubMed: 32267547]
20. Tonon M, Piano S, Gambino CG, et al. Outcomes and Mortality of Grade 1 Ascites and Recurrent Ascites in Patients With Cirrhosis. *Clinical Gastroenterology and Hepatology* 2021;19:358–366. e8. [PubMed: 32272250]
21. Mahmud N, Fricker Z, Serper M, et al. In-Hospital mortality varies by procedure type among cirrhosis surgery admissions. *Liver International* 2019;39:1394–1399. [PubMed: 31141306]
22. Hoteit MA, Ghazale AH, Bain AJ, et al. Model for end-stage liver disease score versus Child score in predicting the outcome of surgical procedures in patients with cirrhosis. *World journal of gastroenterology: WJG* 2008;14:1774. [PubMed: 18350609]
23. Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Annals of internal medicine* 2015;162:W1–W73. [PubMed: 25560730]
24. Nyberg EM, Batech M, Cheetham TC, et al. Postoperative risk of hepatic decompensation after orthopedic surgery in patients with cirrhosis. *Journal of clinical and translational hepatology* 2016;4:83. [PubMed: 27350938]

**What You Need to Know:****Background:**

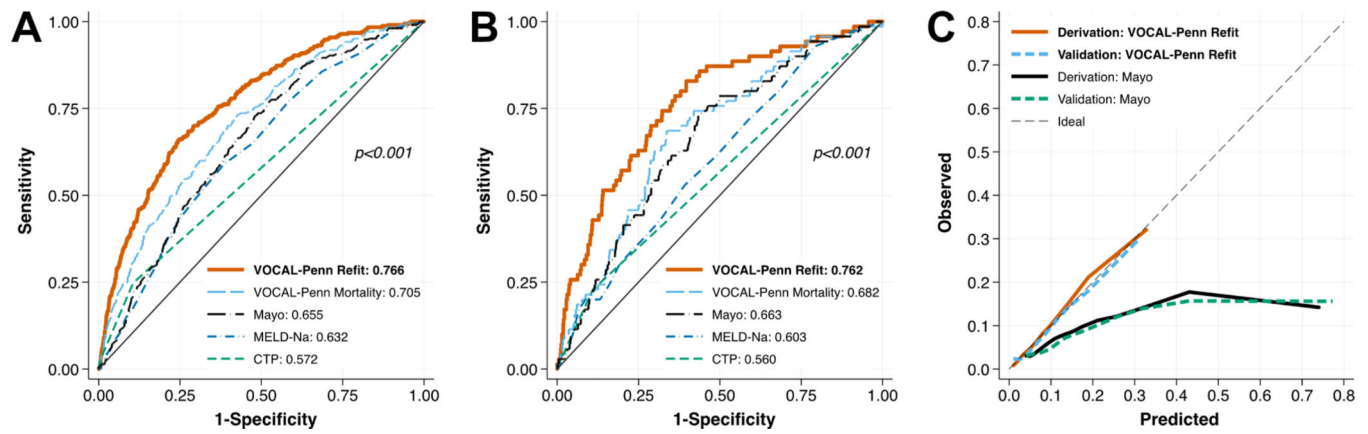
Patients with cirrhosis have increased post-operative risk, however no prediction tools exist to estimate risk of interval decompensation.

**Findings:**

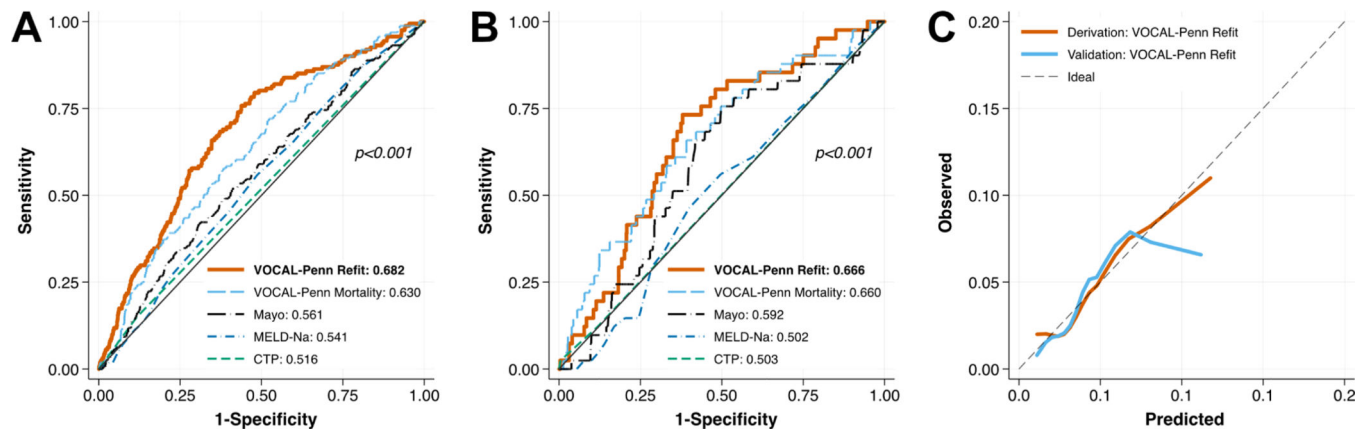
We have developed a novel prediction model (VOCAL-Penn) that accurately predicts 90-day post-operative cirrhosis decompensation, and outperforms other liver disease risk scores.

**Implications for Patient Care:**

Improved prognostication around major surgeries in patients with cirrhosis may help inform risk stratification and counseling for patients and providers.



**Figure 1 –.**  
Receiver Operating Characteristic Curves for 90-Day Post-Operative Decompensation in (A) Derivation and (B) Validation Cohorts, with (C) Associated Calibration Curves



**Figure 2 —.**  
Receiver Operating Characteristic Curves for 90-Day Post-Operative Infection in (A) Derivation and (B) Validation Cohorts, with (C) Associated Calibration Curves



**Table 1:**

## Pre-Operative Cohort Characteristics

Variable	Overall (N = 4,712)	Derivation (N = 3,770)	Validation (N = 942)
Age, median (IQR)	64 (60, 69)	64 (60, 69)	64 (60, 69)
Male Sex	4582 (97.2%)	3670 (97.3%)	912 (96.8%)
Race			
White	2981 (63.3%)	2386 (63.3%)	595 (63.2%)
Black	741 (15.7%)	586 (15.5%)	155 (16.5%)
Hispanic	323 (6.9%)	262 (6.9%)	61 (6.5%)
Asian	51 (1.1%)	47 (1.2%)	4 (0.4%)
Other	616 (13.1%)	489 (13.0%)	127 (13.5%)
Smoking History			
Never smoker	846 (18.8%)	682 (18.1%)	164 (17.4%)
Former smoker	979 (21.8%)	779 (20.7%)	200 (21.2%)
Current smoker	2676 (59.5%)	2147 (56.9%)	529 (56.2%)
Surgery Category			
Abdominal – Laparoscopic	476 (10.1%)	402 (10.7%)	74 (7.9%)
Abdominal – Open	665 (14.1%)	526 (14.0%)	139 (14.8%)
Abdominal Wall	1308 (27.8%)	1033 (27.4%)	275 (29.2%)
Vascular	550 (11.7%)	436 (11.6%)	114 (12.1%)
Major Orthopedic	1298 (27.5%)	1043 (27.7%)	255 (27.1%)
Chest/Cardiac	415 (8.8%)	330 (8.8%)	85 (9.0%)
Emergency Surgery	476 (10.1%)	383 (10.2%)	93 (9.9%)
ASA Classification			
2	200 (4.2%)	161 (4.3%)	39 (4.1%)
3	3196 (67.8%)	2556 (67.8%)	640 (67.9%)
4	1316 (27.9%)	1053 (27.9%)	263 (27.9%)
Sodium, median (IQR)	138 (136, 140)	138 (136, 140)	138 (135, 140)
Creatinine, median (IQR)	1.0 (0.8, 1.2)	.98 (.8, 1.2)	.95 (.8, 1.2)
AST, median (IQR)	32 (22.5, 50)	32 (22.5, 50)	32 (23, 50)
ALT, median (IQR)	27 (18, 43)	27 (18, 42)	27 (18, 45)
Total Bilirubin, median (IQR)	0.8 (0.5, 1.1)	0.8 (0.5, 1.2)	0.7 (0.5, 1.1)
Albumin, median (IQR)	3.7 (3.2, 4.1)	3.7 (3.2, 4.1)	3.7 (3.2, 4.05)
Platelet Count, median (IQR)	152 (107, 207)	153 (107, 207)	150 (109, 207.5)
INR, median (IQR)	1.1 (1.0, 1.2)	1.1 (1.03, 1.25)	1.1 (1.01, 1.22)
MELD, median (IQR)	8 (7, 11)	9 (7, 11)	8 (7, 11)
MELD-Na, median (IQR)	10 (8, 14)	10 (8, 14)	10 (8, 14)
Child-Turcotte-Pugh Class			
A	4159 (88.3%)	3332 (88.4%)	827 (87.8%)

Variable	Overall (N = 4,712)	Derivation (N = 3,770)	Validation (N = 942)
B	530 (11.2%)	420 (11.1%)	110 (11.7%)
C	23 (0.5%)	18 (0.5%)	5 (0.5%)
<b>Ascites Category</b>			
None	4092 (86.8%)	3262 (86.5%)	830 (88.1%)
Slight	469 (10.0%)	389 (10.3%)	80 (8.5%)
Moderate	151 (3.2%)	119 (3.2%)	32 (3.4%)
<b>Hepatic Encephalopathy</b>			
No encephalopathy	4577 (97.1%)	3658 (97.0%)	919 (97.6%)
Grade 1–2	122 (2.6%)	102 (2.7%)	20 (2.1%)
Grade 3–4	13 (0.3%)	10 (0.3%)	3 (0.3%)
<b>Etiology of Liver Disease</b>			
Hepatitis C	612 (13.0%)	503 (13.3%)	109 (11.6%)
Hepatitis B	73 (1.5%)	53 (1.4%)	20 (2.1%)
Alcohol-related liver disease	1662 (35.3%)	1322 (35.1%)	340 (36.1%)
Hepatitis C + Alcohol	1388 (29.5%)	1101 (29.2%)	287 (30.5%)
Fatty Liver Disease	585 (12.4%)	482 (12.8%)	103 (10.9%)
Other	392 (8.3%)	309 (8.2%)	83 (8.8%)
<b>History of Prior Decompensation</b>	2066 (43.8%)	1654 (43.9%)	412 (43.7%)
<b>Prior TIPS</b>	63 (1.3%)	49 (1.3%)	14 (1.5%)
<b>Hypertension</b>	3904 (85.4%)	3146 (85.8%)	758 (83.9%)
<b>Diabetes</b>	2355 (51.7%)	1907 (52.2%)	448 (49.7%)
<b>Obesity (body mass index ≥ 30)</b>	3359 (73.8%)	2675 (73.3%)	684 (75.9%)
<b>Atrial Fibrillation</b>	803 (17.0%)	625 (16.6%)	178 (18.9%)
<b>Coronary Artery Disease</b>	1622 (34.4%)	1298 (34.4%)	324 (34.4%)
<b>Congestive Heart Failure</b>	1239 (26.3%)	986 (26.2%)	253 (26.9%)

Abbreviations: IQR = interquartile range; ASA = American Society of Anesthesiologists; TIPS = transjugular intrahepatic portosystemic shunt

**Table 2 –**

Decompensation and Infection through 90 and 30 Post-operative Days, Respectively, Stratified by Surgery Category

<b>Surgery Category</b>	<b>90-day Decompensation</b>	<b>30-day Infection</b>
Abdominal – Laparoscopic (N = 476)	35 (7.4%)	22 (4.6%)
Abdominal – Open (N = 665)	119 (17.9%)	70 (10.5%)
Abdominal Wall (N = 1308)	105 (8.0%)	37 (2.8%)
Vascular (N = 550)	26 (4.7%)	33 (6.0%)
Major Orthopedic (N = 1298)	86 (6.6%)	62 (4.8%)
Chest/Cardiac (N = 415)	37 (8.9%)	20 (4.8%)
<i>P-value</i>	<i>&lt;0.001</i>	<i>&lt;0.001</i>

**Table 3 –**

Multivariable Logistic Regression Models for Decompensation and Infection through 90 and 30 Post-operative Days, Respectively

Variable	Odds Ratio	95% Confidence Interval	P-value
<b>90-day Post-Operative Decompensation Model</b>			
<b>ASA Class (Ordinal; 2, 3, or 4)</b>	1.36	(1.05 – 1.76)	0.02
<b>Emergency Procedure</b>	1.66	(1.19 – 2.32)	0.003
<b>Age (per year)</b>	1.01	(0.99 – 1.02)	0.44
<b>Surgery Category</b>			
Abdominal – Laparoscopic	1.0 (Ref)		
Abdominal – Open	1.84	(1.12 – 3.02)	0.02
Abdominal Wall	0.86	(0.53 – 1.38)	0.53
Vascular	0.62	(0.34 – 1.16)	0.13
Major Orthopedic	0.71	(0.43 – 1.16)	0.17
Chest/Cardiac	0.84	(0.46 – 1.55)	0.58
<b>Platelet Count (per 1,000/<math>\mu</math>L)</b>			
Spline 1	0.99	(0.99 – 1.00)	0.007
Spline 2	1.00	(1.00 – 1.01)	0.41
<b>Albumin (per 1 g/dL)</b>			
Spline 1	0.64	(0.46 – 0.88)	0.006
Spline 2	0.63	(0.38 – 1.05)	0.08
<b>Total Bilirubin (per 1 mg/dL)</b>			
Spline 1	4.21	(2.11 – 8.41)	<0.001
Spline 2	0.15	(0.06 – 0.38)	<0.001
<b>Obesity (body mass index <math>\geq 30</math>)</b>	0.84	(0.63 – 1.11)	0.21
<b>Etiology of Cirrhosis</b>			
Non-NAFLD	1.0 (Ref)		
NAFLD	1.28	(0.90 – 1.83)	0.17
<b>30-day Post-Operative Infection Model</b>			
<b>ASA Class (Ordinal; 2, 3, or 4)</b>	0.94	(0.68 – 1.29)	0.69
<b>Emergency Procedure</b>	1.35	(0.87 – 2.11)	0.18
<b>Age (per year)</b>	1.01	(0.99 – 1.03)	0.26
<b>Surgery Category</b>			
Abdominal – Laparoscopic	1.0 (Ref)		
Abdominal – Open	2.16	(1.18 – 3.95)	0.01*
Abdominal Wall	0.60	(0.31 – 1.13)	0.11
Vascular	1.49	(0.77 – 2.87)	0.24
Major Orthopedic	0.98	(0.54 – 1.77)	0.93
Chest/Cardiac	1.01	(0.47 – 2.15)	0.99
<b>Platelet Count (per 1,000/<math>\mu</math>L)</b>			

Variable	Odds Ratio	95% Confidence Interval	P-value
Spline 1	1.00	(1.00 – 1.01)	0.21
Spline 2	1.00	(0.99 – 1.00)	0.23
<b>Albumin (per 1 g/dL)</b>			
Spline 1	0.71	(0.47 – 1.06)	0.09
Spline 2	0.71	(0.40 – 1.26)	0.24
<b>Total Bilirubin (per 1 mg/dL)</b>			
Spline 1	1.19	(0.56 – 2.54)	0.65
Spline 2	0.85	(0.30 – 2.38)	0.76
<b>Obesity (body mass index ≥ 30)</b>	0.84	(0.60 – 1.19)	0.33
<b>Etiology of Cirrhosis</b>			
Non-NAFLD	1.0 (Ref)		
NAFLD	1.12	(0.72 – 1.73)	0.62

Abbreviations: ASA = American Society of Anesthesiologists; NAFLD = non-alcoholic fatty liver disease

**Table 4 –**

Prediction Model Discrimination (C-statistics) for 90-Day Post-Operative Decompensation and 30-Day Post-Operative Infection

		90-Day Decompensation		30-Day Infection	
	Risk Score	C-statistic (95% CI)	P value*	C-statistic (95% CI)	P value*
<b>Derivation</b>	<b>VOCAL-Penn Refit</b>	<b>0.766 (0.738 – 0.793)</b>	<b>-</b>	<b>0.692 (0.653 – 0.730)</b>	<b>-</b>
	VOCAL-Penn Mortality	0.705 (0.676 – 0.735)	<0.001	0.623 (0.583 – 0.664)	<0.001
	Mayo Score	0.655 (0.626 – 0.684)	<0.001	0.563 (0.520 – 0.606)	<0.001
	MELD-Sodium	0.632 (0.601 – 0.663)	<0.001	0.546 (0.504 – 0.588)	<0.001
	CTP	0.572 (0.547 – 0.596)	<0.001	0.529 (0.500 – 0.557)	<0.001
<b>Validation</b>	<b>VOCAL-Penn Refit</b>	<b>0.762 (0.704 – 0.821)</b>	<b>-</b>	<b>0.654 (0.577 – 0.730)</b>	<b>-</b>
	VOCAL-Penn Mortality	0.682 (0.619 – 0.745)	<0.001	0.634 (0.549 – 0.718)	0.49
	Mayo Score	0.663 (0.600 – 0.727)	0.005	0.603 (0.530 – 0.676)	0.30
	MELD-Sodium	0.603 (0.536 – 0.670)	<0.001	0.561 (0.484 – 0.638)	0.046
	CTP	0.560 (0.509 – 0.611)	<0.001	0.529 (0.476 – 0.583)	0.002

Abbreviations: CI = confidence interval; MELD = model for end-stage liver disease; CTP = Child-Turcotte-Pugh

\* P values correspond to pairwise comparisons between VOCAL-Penn Refit and other scores. A Bonferroni adjusted threshold of alpha = 1.25% was used to determine statistical significance.