Quick chronic liver failure-sequential organ failure assessment: an easy-to-use scoring model for predicting mortality risk in critically ill cirrhosis patients

Xiao-Dong Zhou, Jia-Ying Zhang, Wen-Yue Liu, Sheng-Jie Wu, Ke-Qing Shi, Martin Braddock, Yong-Ping Chen, Wei-Jian Huang and Ming-Hua Zheng

**Background and aim** Critically ill cirrhosis patients have an increased risk of morbidity and mortality, even after admission to the ICU. Our objectives were to compare the predictive accuracy of model for end-stage liver disease (MELD), MELD-Na, UK model for end-stage liver disease, and chronic liver failure-sequential organ failure assessment (CLIF-SOFA) by the development and validation of an easy-to-use prognostic model [named quick CLIF-SOFA (qCLIF-SOFA)] for early risk prediction in critically ill patients with cirrhosis.

**Patients and methods** Overall, 1460 patients were extracted from the MIMIC-III database and enrolled in this study at 30-day and 90-day follow-up. qCLIF-SOFA was developed in the established cohort (n = 730) and a performance analysis was completed in the validation cohort (n = 730) using area under the receiver operating characteristic curve. Results were compared with CLIF-SOFA.

**Results** The performance of CLIF-SOFA was significantly better than that of MELD, MELD-Na, and UK model for end-stage liver disease for predicting both 30-day and 90-day mortality (all P < 0.05). qCLIF-SOFA consisted of five independent factors (bilirubin, creatinine, international normalized ratio, mean arterial pressure, and vasopressin) associated with mortality. In the established cohort, CLIF-SOFA and qCLIF-SOFA predicted mortality with area under the receiver operating characteristic curve values of 0.768 versus 0.743 at 30-day, 0.747 versus 0.744 at 90-day, and 0.699 versus 0.706 at 1 year, respectively (all P > 0.05). A similar result was observed in the validation cohort (0.735 vs. 0.734 at 30 days, 0.723 vs. 0.737 at 90 days, and 0.682 vs. 0.700 at 1 year, respectively, all P > 0.05).

**Conclusion** The utility of CLIF-SOFA was further shown to predict mortality for critically ill cirrhosis patients. The novel and simpler qCLIF-SOFA model showed comparable accuracy compared with existing CLIF-SOFA for prognostic prediction. Eur J Gastroenterol Hepatol 00:000–000

Copyright © 2017 Wolters Kluwer Health, Inc. All rights reserved.

**Introduction** Critically ill cirrhosis patients (CICP), despite admission to the ICU and aggressive interventions, show rapid disease progression and high short-term mortality [1–4]. Severity assessment models are urgently needed to risk stratify patients to determine an optimum therapeutic approach and regimen and to ensure appropriate allocation of medical resources [5–8].

Chronic liver failure-sequential organ failure assessment (CLIF-SOFA) score, as a modified SOFA score, is an excellent prognostic evaluation tool for CICP [3,9]. Models for end-stage liver disease (MELD), MELD-Na, and the UK model for end-stage liver disease (UKELD) are also widely utilized for evaluating the severity of CICP [10–13]. In addition, considerable effort has focused on determining whether CLIF-SOFA may outperform MELD-based scoring systems for predicting short-term and long-term mortality of CICP [8,9,14]. Pan et al. [9] showed that CLIF-SOFA appeared to be superior to MELD in predicting mortality risk, whereas Jalan and colleagues showed that CLIF-SOFA does not significantly improve the accuracy of prediction of MELD and MELD-Na models [14]. Several scoring systems modified by CLIF-SOFA were developed and validated with laboratory or demographic parameters and showed an improved accuracy of prediction [6,15]. However, CLIF-SOFA and the modified CLIF-SOFA scores were relatively complex to generate and interpret, and the multiple variables required
for the scoring to algorithms were not convenient or simple for clinicians to use.

The SOFA score included seven clinical variables (range: 0–24 points) and the complex calculation restricted its use for early risk prediction especially in the absence of some relevant variables [12,16]. Seymour et al. [17] derived and validated the quick SOFA model for predicting mortality for sepsis in patients outside of the ICU and confirmed its predictive performance. Encouraged by these findings, we aimed to create a quick chronic liver failure-sequential organ failure assessment (qCLIF-SOFA) score on the basis of original CLIF-SOFA variables to improve the clinical utility of CLIF-SOFA for early risk prediction in CCIP.

In this study, the objectives were as follows: (i) to evaluate the performance of CLIF-SOFA compared with that of MELD, MELD-Na, and UKELD for predicting short-term mortality and long-term mortality; (ii) to derive and examine the predictive efficiency of the easy-to-use model (qCLIF-SOFA) compared with CLIF-SOFA for mortality risk prediction for CICP.

Patients and methods

Study design

Patient data were derived from the Multi-parameter Intelligent Monitoring in Intensive Care database (version 2.6). This is a publicly available large-scale ICU database, currently consisting of more than 40 000 ICU patients admitted to Beth Israel Deaconess Medical Center (Boston, Massachusetts, USA) from July 2001 to June 2008. A total of 1460 CICP were extracted and enrolled in this study. Our permission to access the database was approved after competition of the National Institute of Health web-based training course named ‘Protecting Human Research Participants’ (our certification number: 1605699).

Definition

CICP was defined as patients with cirrhosis admitted to ICU. Liver cirrhosis was defined when at least two of the following criteria were fulfilled: (i) ultrasonographic evidence of a small-sized liver and with or without splenomegaly/ascites; (ii) hypoalbuminemia (serum albumin < 3.5 g/l); and (iii) more than two of aminotransferase to platelet ratio (×10⁷/l) × 100. Alcoholic cirrhosis of liver was established as a daily alcohol consumption of more than 80 g/day for at least 5 years.

Data collection

Our investigators extracted detailed patient records for CICP, typically containing 24 h of continuous real-time signals and laboratory parameters, together with other clinical data. The continuous real-time signals, which included heart rate, respiration, temperature, systolic blood pressure, diastolic blood pressure, and mean arterial pressure (MAP), were derived by ICU nurses from the hospital’s on-line information systems. Laboratory parameters from routine tests, including glucose, white blood cell, platelet, sodium, potassium, blood urea nitrogen, PaO₂, PCO₂, FIO₂, bicarbonate, creatinine, lactate, international normalized ratio (INR), and bilirubin, were organized into a relational database. The other clinical data included age, sex, height, weight, ethnicity, vasopressin use, liver cirrhosis, and survival time. Mortality data were censored after hospital discharge and were obtained by Social Security Death Records from the US government. All patients were required to be scored using CLIF-SOFA, MELD, MELD-Na, and UKELD models using the variables described previously. The start date was the date of the patient’s admission and the primary end points were defined at 30-day, 90-day, and 1-year all-cause mortality.

Data extraction was performed using Oracle SQL Developer, version 3.0 (Oracle Corporation, Redwood Shores, California, USA). As this was a retrospective study, no ethical approval was required for these analyses of nonpatient identifiable and anonymous database.

CLIF-SOFA, MELD, MELD-Na, and UKELD

CLIF-SOFA was calculated according to published formulae [12]. MELD, MELD-Na and UKELD were computed using previously described variables as follows: MELD: \( R = 9.57 \times \log_e [\text{creatinine (mg/dl)}] + 3.78 \times \log_e [\text{bilirubin (mg/dl)}] + 11.2 \times \log_e (\text{INR}) + 6.43 \) [18]. In addition, MELD-Na: \( R = \text{MELD} + 1.59 \times [135 - \text{Na (mmol/l)}] \) [11]; UKELD: \( 5 \times [1.5 \times \log_e (\text{INR}) + 0.3 \times \log_e [\text{creatinine (µmol/l)}] + 0.6 \times \log_e [\text{bilirubin (µmol/l)}] - 13 \times \log_e [\text{Na (mmol/l) + 70}]] \) [13].

Statistical analysis

Data were presented as mean with SDs for continuous and normally distributed variables, or as frequency (%) for categorical variables. The Kolmogorov–Smirnov test was carried out to check for the distribution of variables. For comparisons, Student’s t-test and the Mann–Whitney test were used for continuous baseline characteristics of the each group for continuous variables with or without a normal distribution, respectively. The χ²-test was performed for categorical variables. Univariate logistic analyses were carried out to determine the unadjusted association of clinical and laboratory parameters with prognosis. All clinical and statistical parameters (P < 0.10), especially from the original CLIF-SOFA, were included as candidate variables in a forward conditional stepwise logistic regression analysis to identify the final variables. Multiple logistic regression models were performed for mortality prediction. Survival curves were constructed based on Kaplan–Meier estimates and comparisons were performed using the log-rank test.

All patients were enrolled for comparison of the discriminative value among CLIF-SOFA MELD, MELD-Na, and UKELD models for predicting mortality risks at 30 and 90 days. To develop and validate qCLIF-SOFA, patients were stratified into two cohorts according to the distribution of the year of admission (the established cohort: from July 2004 to June 2008, the validation cohort: from July 2001 to June 2004). qCLIF-SOFA was developed in the established cohort and a performance analysis was completed in the validation cohort using area under the receiver operating curve (AUROC).

Predictions of 30-day, 90-day, and 1-year mortality by various scores in addition to qCLIF-SOFA were assessed using AUROC. The optimal cutoff point was identified with
the maximal Youden index (sensitivity + specificity – 1). In addition, the corresponding sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, positive predictive value, and negative predictive value were calculated according to the AUROC results. The reporting followed the recently published Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis statement [19].

A two-tailed P-value of less than 0.05 was considered statistically significant. All P values calculated were two-tailed and significance was defined at the 95% level. Statistical analyses were carried out using SPSS, version 23.0 software (IBM, Armonk, New York, USA), MedCalc, version 12.7 (MedCalc Software, Ostend, Belgium).

**Results**

**Baseline characteristics of the study population**

A total of 1460 consecutive CICP were identified, of whom 730 were admitted in the established cohort during the period from July 2004 to June 2008 (Table 1). In the established cohort, the mean age of the patients was 57.6 ± 11.8 years and there were 506 (69.3%) men. Alcoholic cirrhosis of liver was the most common etiology of cirrhosis (49.0%). Moreover, the most common ethnicity of CICP was White (69.9%). The validation cohort included 730 patients during the period from July 2001 to June 2004. The mean age was 57.6 ± 11.4 years and men were predominant (67.9%), of whom 49.2% of patients (n = 359) were diagnosed with alcoholic cirrhosis of the liver and 1.6% of patients (n = 12) were diagnosed with biliary cirrhosis. The established cohort had a similar 30-day, 90-day, and 1-year mortality (29.2, 38.4, and 47.8%, respectively) as patients in the validation cohort (26.8, 36, and 46%, respectively), with no statistically significant difference (P = 0.351, 0.386, and 0.529, respectively). Similar demographic data and clinical characteristics of both survival and mortality groups are shown in Table 2. Compared with patients in the survival group, patients who died were older and had significantly higher serum bilirubin, INR, serum creatinine, and vasopressin used ratio (P < 0.05). The distributions of patients over CLIF-SOFA (range: 0–21) and qCLIF-SOFA (range: 0–5) are listed in Supplementary Fig. 1 (Supplemental digital content 1, http://links.lww.com/EJGH/A178). In the established cohort, 43.3% of patients (n = 316) had two or more qCLIF-SOFA values and the validation cohort had a similar population distribution [302 (41.3%)].

**Performance of CLIF-SOFA compared with MELD, MELD-Na, and UKELD**

Figure 1 shows the discriminatory value of CLIF-SOFA to predict 30-day mortality risk and 90-day mortality in CICP. As shown in Table 2, MELD had an AUROC of 0.725 [95% confidence interval (CI): 0.691–0.757], MELD-Na of 0.676 [95% CI: 0.641–0.710], and UKELD of 0.578 [95% CI: 0.542–0.615]. In the same dataset, CLIF-SOFA with an AUROC of 0.768 (95% CI: 0.736–0.799) showed statistical improvement compared with MELD, MELD-Na, and UKELD in predicting mortality. When predicting 90-day mortality, the AUROC for CLIF-SOFA, MELD, MELD-Na, and UKELD were as follows: CLIF-SOFA 0.747 (95% CI: 0.714–0.778), MELD 0.712 (95% CI: 0.677–0.744), MELD-Na 0.711 (95% CI: 0.677–0.744), and UKELD 0.703 (95% CI: 0.668–0.736). Therefore, CLIF-SOFA showed a significantly higher discriminative value than the MELD, MELD-Na, and UKELD at 30 and 90 days (all P < 0.05) (Table 3).

![Image](http://links.lww.com/EJGH/A178)
Development of qCLIF-SOFA

Multiple logistic analysis showed that bilirubin (1.037, 95% CI: 1.014–1.061, P < 0.001), creatinine (1.248, 95% CI: 1.107–1.407, P < 0.001), INR (1.426, 95% CI: 1.134–1.794, P < 0.002), MAP (0.978, 95% CI: 0.968–0.988, P < 0.001), and vasopressin (1.984, 95% CI: 1.432–2.749, P < 0.001) were identified as independent risk factors for the mortality of CICP (Table 4).

Five optimal cut-points were selected to distinguish two categorical variables that were associated directly with mortality risk, which maximized the discriminatory value of qCLIF-SOFA (range: 0–5) to predict mortality risk at each of the study time points.

Subgroup analysis showed that patients with bilirubin of 6.2 mg/dl or more or creatinine of 1.9 mg/dl or more or INR 1.9 or more or MAP of 70 mmHg or less with vasopressin use had a poorer overall survival compared with patients who did not have these characteristics (Supplementary Fig. 2, Supplemental digital content 2, http://links.lww.com/EJGH/A179).

Construction of qCLIF-SOFA

The final qCLIF-SOFA model included bilirubin of 6.2 mg/dl or more, creatinine of 1.9 mg/dl or more, INR of 1.9 or more, a mean arterial pressure of 70 mmHg or less, and vasopressin use (yes) (1 point each; score range: 0–5) (Table 5).

Performance of qCLIF-SOFA

The performance analysis result is presented in Fig. 2 and Table 6. The AUROC of qCLIF-SOFA was 0.768 (95% CI: 0.742–0.794) in both the validation and established cohort, respectively.
0.736–0.799) in predicting 30-day mortality, 0.747 (95% CI: 0.714–0.778), 90-day mortality, and 0.699 (95% CI: 0.665–0.732) for 1-year mortality in the established cohort. CLIF-SOFA appears to have better discriminative capabilities than those of qCLIF-SOFA, but shows no statistical significance (P = 0.121 at 30 days, P = 0.832 at 90 days, P = 0.661 at 1 year). When applying the optimal cutoff value of 1.5 for qCLIF-SOFA in predicting 30-day mortality, the sensitivity and specificity were 69.48 and 67.50%, respectively. Moreover, the positive likelihood ratio was 2.14, the negative likelihood ratio was 0.45, and the positive predictive and negative predictive values were 46.8 and 84.3, respectively.

In the validation cohort, the AUROC of CLIF-SOFA and qCLIF-SOFA was 0.735 (95% CI: 0.701–0.766) and 0.734 (95% CI: 0.700–0.765). No statistical difference was observed between CLIF-SOFA and qCLIF-SOFA at each of the study time points (P = 0.946 at 30 days, P = 0.358 at 90 days, P = 0.219 at 1 year).

Table 3. Diagnostic accuracy of CLIF-SOFA, MELD, MELD-Na, and UKELD in the established cohort and validation cohort at cutoff points and at different time periods

<table>
<thead>
<tr>
<th>Models</th>
<th>30 Days</th>
<th>P-value*</th>
<th>90 Days</th>
<th>P-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLIF-SOFA</td>
<td>0.768 (0.736–0.799)</td>
<td>–</td>
<td>0.747 (0.714–0.778)</td>
<td>–</td>
</tr>
<tr>
<td>MELD</td>
<td>0.725 (0.691–0.757)</td>
<td>0.012</td>
<td>0.712 (0.677–0.744)</td>
<td>0.035</td>
</tr>
<tr>
<td>MELD-Na</td>
<td>0.676 (0.641–0.710)</td>
<td>&lt; 0.001</td>
<td>0.711 (0.677–0.744)</td>
<td>0.034</td>
</tr>
<tr>
<td>UKELD</td>
<td>0.578 (0.542–0.615)</td>
<td>&lt; 0.001</td>
<td>0.703 (0.668–0.738)</td>
<td>0.014</td>
</tr>
</tbody>
</table>

Table 4. Multivariate analysis of association between mortality and clinical/laboratory characteristics in critically ill cirrhosis patients in the established cohort

<table>
<thead>
<tr>
<th>Variables</th>
<th>B</th>
<th>HR</th>
<th>95% CI</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bilirubin</td>
<td>0.036</td>
<td>1.037</td>
<td>1.014–1.061</td>
<td>0.001</td>
</tr>
<tr>
<td>Creatinine</td>
<td>0.221</td>
<td>1.248</td>
<td>1.107–1.407</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>INR</td>
<td>0.355</td>
<td>1.426</td>
<td>1.134–1.794</td>
<td>0.002</td>
</tr>
<tr>
<td>MAP</td>
<td>−0.022</td>
<td>0.978</td>
<td>0.968–0.988</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Vasopressin</td>
<td>0.685</td>
<td>1.984</td>
<td>1.432–2.749</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

Discussion

Several previous studies aimed to determine whether CLIF-SOFA is the optimal prognosis for CICP [6,10,18,20]. In this present study, the predictive performance of CLIF-SOFA was confirmed to be superior to that of MELD, MELD-Na, and UKELD in predicting mortality for CICP.

Many risk prognostic models, related to critical ill cirrhosis, have been developed and validated previously [8,15,21–24]. Zauner et al. [22] developed the intensive care cirrhosis outcome score from 196 CICP, using...
multiple logistic regression analysis, with four variables include bilirubin, cholesterol, creatinine clearance, and lactate. The royal free hospital score was derived from 635 CICP between 1989 and 2012, with arterial lactate, total bilirubin, INR, variceal bleeding, alveolar-arterial gradient, and urea [4,22]. To avoid using complex formulae, Edmark et al. [23] generated the LiFe (liver, injury, failure, evaluation) score from 157 CICP in the European Society for Intensive Care Medicine guidelines, with only three laboratory-based variables (arterial lactate, total bilirubin, and INR, each for 0–3 points). Jalan et al. [8] had simplified the original CLIF-SOFA for combining age and log-transformed white blood cell count to improve the predictive accuracy of mortality in acute-on-chronic liver failure patients, but not in CICP.
This research is the first large-scale study to generate and validate an easy-to-use clinical risk prediction score (qCLIF-SOFA) based on original CLIF-SOFA for predicting mortality in CICP. qCLIF-SOFA scoring model consists of five variables: bilirubin, INR, MAP, creatinine, and vasopressin used (0 or 1 for each variable, range: 0–5 points). Elevated bilirubin concentration and INR level has been shown to be associated with mortality of patients with liver disease [25–27]. Patients with elevated serum creatinine can suffer from more serious kidney injury, increasing their risk of mortality [28–31]. The decrease in MAP or vasopressin use always occurred with organ deterioration and failure associated with mortality [32,33].

Despite simplifying the original CLIF-SOFA model, the predictive performance of qCLIF-SOFA was not inferior to CLIF-SOFA and remains statistically significant. However, although qCLIF-SOFA was simpler to calculate from variables that could be obtained from case histories, it did not add to the discriminative accuracy of CLIF-SOFA. CLIF-SOFA still plays an irreplaceable role in assessing the severity of organ failures and predicting the mortality of CICP.

A few potential limitations need consideration. First, because our study population was exclusively from a single center, potential selection bias might exist. Second, in the absence of a comparative large-scale population, the performance analysis of our final predictive model had to be completed on the same database. Multicenter large-scale studies for CICP with more than 1 year of follow-up are needed to further verify its applicability of qCLIF-SOFA. Third, we did not perform a comparison between qCLIF-SOFA and other published modified CLIF-SOFA scoring models.

**Conclusion**

CLIF-SOFA was confirmed to show discriminability in predicting mortality for CICP. The novel qCLIF-SOFA, as a simpler bedside model, showed a comparable accuracy compared with existing CLIF-SOFA for prognostic prediction.

**Acknowledgements**

This work was supported by grants from the National Natural Science Foundation of China (81500665), Scientific Research Foundation of Wenzhou (Y20160223), the High Level Creative Talents from Department of Public Health in Zhejiang Province, and the Project of New Century 551 Talent Nurturing in Wenzhou. Zhou X.D., Zhang J.Y., Huang W.J., and Zheng M.H. designed the study. Zhou X.D., Zhang J.Y., and Liu W.Y. collected data. Shi K.Q., Zhou X.D. carried out the statistical analyses. Liu W.Y., Wu S.J. prepared figures. Zhou X.D., Zhang J.Y., Braddock M., Chen Y.P., Huang W.J., and Zheng M.H. reviewed the results, interpreted data, and wrote the manuscript. All authors have made an intellectual contribution to the manuscript and approved the submission.

**Conflicts of interest**

There are no conflicts of interest.

**References**


