Retrospective Cohort Study

Clinical value of predictive model based on liver stiffness measurement in predicting liver reserve function of compensated chronic liver disease

Lai RM et al. The model predicted liver reserve function

Rui-Min Lai, Miao-Miao Wang, Xiao-Yu Lin, Qi Zheng, Jing Chen

Abstract

BACKGROUND
Assessment of liver reserve function (LRF) is essential to predict prognosis for patients with chronic liver disease (CLD) and determines the extent of liver resection patients with hepatocellular carcinoma.

AIM
To establish noninvasive models of LRF assessment based on liver stiffness measurement (LSM) and to evaluate its clinical performance.

METHODS
A total of 360 patients with compensated CLD were retrospectively analyzed as a training cohort. The new predictive models were established through logistic regression analysis and were validated internally in a prospective cohort (132 patients).

RESULTS
Our study defined indocyanine green retention rate at 15 min (ICGR15) ≥ 10% as mildly impaired LRF and ICGR15 ≥ 20% as severely impaired LRF. We constructed predictive models of LRF, called the mLPaM and sLPaM, which involved only LSM, prothrombin time international normalized ratio to albumin ratio (PTAR), age and end-stage liver disease (MELD). The Area under the curve of mLPaM model (0.855, 0.872, respectively) and sLPaM model (0.869, 0.876, respectively) were higher than that of methods for MELD, albumin-bilirubin grade and PTAR in the two cohorts, and their sensitivity and negative predictive value were the highest among these methods in the training cohort. Meanwhile, the new models showed good sensitivity and accuracy for LRF impairment diagnosis in the validation cohort.

CONCLUSION
The new models had a good predictive performance on LRF and could replace indocyanine green (ICG) clearance test, especially for the patients who were unable to undergo ICG testing.

Key Words: Liver stiffness measurement; Chronic liver disease; Liver reserve function; Indocyanine green clearance test; Predictive model


Core Tip: This study aimed to establish a predictive model of liver stiffness measurement (LSM) in patients with the compensated chronic liver disease based on LSM and evaluate its clinical value. The result from our study showed that the new model had a good predictive performance of liver reserve function (LRF). The area under the curve of the model was higher than that of model for end-stage liver disease, albumin-bilirubin grade and prothrombin time international normalized ratio to
albumin ratio. Moreover, the predictive performance of the new model had been validated in prospective cohort. We believed that the model could replace the indocyanine green (ICG) clearance test to assess LRF, especially for the patient who was unable to undergo ICG testing.

INTRODUCTION

The high prevalence of chronic liver disease (CLD) in China has become a severe public health problem. Cirrhosis, hepatocellular carcinoma (HCC), hepatic encephalopathy and other decompensated complications are the leading causes of mortality in the CLD patients without treatment. Liver reserve function (LRF) is defined as the compensated ability of the liver to maintain normal physiological functions in the face of injury, which mainly depends on the quality and quantity of hepatocytes in the remnant liver\textsuperscript{[1,2]}. There are no obvious clinical symptoms of the CLD patients at an early stage, but their LRF may be impaired. Early evaluation of LRF is of great help for identifying disease progression, timely implementing intervention and appropriately taking treatment strategies in the CLD patients. Several scoring systems, including the Child-Turcotte-Pugh (CTP), model for end-stage liver disease (MELD), albumin-bilirubin (ALBI) and APRI, could be used to evaluate the LRF\textsuperscript{[1,3-5]}. Although the CTP score is widely used to assess LRF, it includes subjective criteria, such as ascites and hepatic encephalopathy. The MELD score is initially used as a standard model to assess the prognosis of patients with decompensated cirrhosis, but its creatinine (Cr) value could be significantly affected by age and gender.

The indocyanine green (ICG) clearance test is commonly used for LRF assessment, which is considered the most valuable evaluation method of LRF. ICGR15 had become a standard dynamic preoperative instrument to evaluate the hepatic functional reserve before liver resection and predicted post-hepatectomy liver failure\textsuperscript{[6,7]}. However, the operation process of ICG clearance test is tedious and requires the technical operator, so most of the tests could only be carried out in major hospitals. In addition, some patients are allergic to ICG, which led to the failure of the test. Due to the impossible
implementation of the ICG clearance test in CLD patients, we need to explore a new method to assess LRF accurately.

Liver stiffness measurement (LSM) is commonly used to evaluate the degree of liver fibrosis, and due to its non-invasiveness, cost-efficient and safety, it has been widely applied in clinical treatment. Previous studies showed that LSM could predict the occurrence of liver failure after HCC resection\textsuperscript{[8-10]}. Therefore, LSM has a certain potential value in evaluating hepatic functional reserve.

The purpose of this study was to analyze the association between LSM and ICGR15 in evaluating LRF. We constructed the predictive model based on LSM and explored its clinical application value in evaluating LRF in compensated CLD patients.

**MATERIALS AND METHODS**

**Research population**

All patients with CLD (≥ 18-year-old) consecutively observed in the inpatient department of Hepatology Research Institute of the First Affiliated Hospital, Fujian Medical University, China, from March 2016 to June 2019 were retrospectively analyzed as a training cohort. From September 2019 to August 2020, patients with CLD were prospectively evaluated to validate the new model. Information about the patient’s demographics, ICG clearance test, laboratory data and Fibro-scan examination was abstracted from the electronic medical record system of the First Affiliated Hospital of Fujian Medical University. Patients with the following conditions will be excluded: (1) decompensated cirrhosis with CTP grade B and C; (2) insufficient data; and (3) complicated with other tumors, or gestation. After exclusion, 492 patients were identified for study inclusion, comprising 389 CHB patients, 35 fatty liver disease patients, 21 autoimmune liver disease patients, 8 hepatitis C virus patients and 39 other aetiologies patients. All enrolled patients were divided into a training cohort (360 patients) and validation cohort (132 patients), including 105 HCC patients who met the diagnostic criteria for guidelines for diagnosis and treatment of primary liver cancer in China (2019 edition)\textsuperscript{[11]}. 

4/16
Clinical and laboratory parameters

The demographic data collected included age and gender. The clinical laboratory information included prothrombin time (PT), international normalized ratio (INR), total bilirubin (TBIL), aspartate aminotransferase, alanine aminotransferase, albumin (ALB), glomerular filtration rate, alkaline phosphatase, gamma-glutamyltransferase, cholinesterase, platelet count, and hemoglobin. The parameters were detected by Olympus AU2700 automatic biochemical analyzer. The calculation of CTP score included five items, namely ALB, TBIL, PT, hepatic encephalopathy and ascites\textsuperscript{[12]}. The CTP classifications were defined as grade A (5-6 points), grade B (7-9 points), and grade C (10-15 points). The MELD score was calculated by the formula $3.78 \times \ln[\text{TBIL (mg/dL)}] + 11.2 \times \ln(\text{INR}) + 9.57 \times \ln(\text{Cr (mg/dL)}) + 6.43 \times \text{etiology (cholestasis and alcohol for 0, others for 1) criteria}$\textsuperscript{[13]}. The prothrombin time international normalized ratio to albumin ratio (PTAR) score was calculated by the formula INR/ALB (g/dL)$^{[14]}$. The ALBI score was calculated by the formula $\ln[\text{TBIL (mol/L)}] \times 0.66 + \ln[\text{ALB (g/L)}] - 0.0852$\textsuperscript{[15]}.

All patients received ICG clearance test after overnight fasting, a dose of 0.5 mg/kg of ICG was rapidly injected into patients via a peripheral vein of the forearm. An optical probe attached to the patient’s nose was used to monitor plasma ICG concentrations, and the value of ICG-R15 was calculated by a Pulse Dye Densito-Graph Analyzer (DDG-3300K, Nihon Kohden, Tokyo, Japan)$^{[16]}$. The LRF was defined that ICGR15 < 10% was normal, ICGR15 ≥ 10% was mild impairment, and ICGR15 ≥ 20% was severe impairment.

The Fibro-Scan 502 Touch (Echosens, Paris, France) was performed by the same trained operator and operated according to the manufacturer’s instructions. LSM was performed on the right lobe of the liver through the intercostal spaces. Ten successful acquisitions were performed for every patient. The success Rate (≥ 60%) was calculated as the number of successful measurements divided by the total number of measurements recorded$^{[17]}$. LSM was expressed as the median and IQR [in kilopascals
(kPa)] of all valid measurements obtained. A LSM was considered reliable if 10 valid acquisitions were obtained. Patients with poorly reliable measurements (IQR/median ratio > 0.30 with a median LSM > 7.1 kPa) were excluded\textsuperscript{[18]}. This retrospective study was approved by the ethics committee at the First Affiliated Hospital of Fujian Medical University, China.

**Statistical analysis**

Statistical analyses were performed using SPSS 23.0. The normally distributed continuous variables were presented as mean ± SD, which were further evaluated by Student's $t$-test in the different groups. Whereas, variables showing skewed distributions were evaluated by the Mann-Whitney $U$ test, which were presented as median (interquartile range). Categorical variables were described using frequencies and proportions, and Pearson's chi-squared test was used to compare categorical variables.

Multivariable analyses were conducted on variables that reached $P < 0.1$ at univariate analysis. Multivariate analysis was performed using the logistic regression analysis, and we established regression prediction models to predict the hepatic functional reserve. The continuous variables (cut-off value of LSM for 12.4 and PTAR for 0.280) were transformed into dichotomous variables. In order to avoid collinearity of some clinical indicators, stepwise forward regression was used in multivariate analysis. The optimal cut-off level of the model was determined by a receiver operator characteristic analysis. The areas under the curve (AUCs) were measured and compared to evaluate the discrimination ability of different models. The final predictive model was fitted on an internal validation dataset and on the entire prospective population. A two sided $P$ value less than 0.05 was considered significant.

**RESULTS**

*Summary of baseline clinical and demographic data of chronic liver disease patients*
Overall 392 patients were included in the study, and 103 patients with HCC (Table 1). 350 (71.14%) out of 492 patients were male, the predominant etiology of the liver disease was related to HBV ($n = 389, 79.07\%$). Patients in the validation cohort were elder than those in the training cohort (mean age, $54.84 \pm 27.70$ vs $48.71 \pm 13.34$, $P < 0.001$), as well as there existed statistically significant differences in ALB and TBIL levels. However, the two cohorts had a similar level of LSM and MELD ($P = 0.066$, $P = 0.241$, respectively).

To construct the LRF predictive model based on LSM

With ICGR15 $\geq 10\%$ and ICGR15 $\geq 20\%$ as the predictive points, the new models of mildly impaired LRF (mLPaM) and severely impaired LRF (sLPaM) were constructed based on LSM. In the training cohort, 360 patient variables were included in the multivariate Logistic stepwise regression analysis. LSM (OR = 4.357, 95% CI: 2.248-8.445), PTAR (OR = 3.544, 95% CI: 1.838-6.835), age (OR = 1.048, 95% CI: 1.024-1.073) and MELD (OR = 1.340, 95% CI: 1.150-1.562) were independent influencing factors of ICGR15 $\geq 10\%$ (Table 2). LSM (OR = 3.120, 95% CI: 1.125-8.656), PTAR (OR = 3.524, 95% CI: 1.267-9.801), age (OR = 1.059, 95% CI: 1.024-1.096) and MELD (OR = 1.377, 95% CI: 1.146-1.655) were independent influencing factors of ICGR15 $\geq 20\%$ (Table 3). The predictive models using the above 4 variables were constructed as follows: mLPaM = 1.472 LSM (LSM $\geq 12.4$ for 2, LSM $< 12.4$ for 1) + 1.265 PTAR (PTAR $\geq 0.280$ for 2, PTAR $< 0.280$ for 1) + 0.047 age (years) + 0.291 MELD-7.600 and sLPaM = 1.138 LSM (LSM $\geq 12.4$ for 2, LSM $< 12.4$ for 1) + 1.260 PTAR (PTAR $\geq 0.280$ for 2, PTAR $< 0.280$ for 1) + 0.058 age (years) + 0.320 MELD-9.750.

A comparison of the predictive performance of the constructed model and other methods in the training cohort

The AUC values of mLPaM model (0.855) and sLPaM model (0.872) were greater than that of MELD score, PTAR and ALBI evaluation tools, and their sensitivity and negative predictive values were the best among these evaluation methods (Table 4 and Figure 1).
Internal validation of the new predictive model in the validation cohort

132 CLD patients were prospectively considered for enrollment in the internal validation cohort. The performance of the various methods at predicting LRF was reported (Table 5). The AUC values of mLPaM model (0.869) and sLPaM model (0.876) were greater than other LRF predictive methods. The mLPaM model showed a good sensitivity (89.1%) and an optimal accuracy (78.94%) for mild LRF impairment diagnosis, and sLPaM model reported an optimal sensitivity (92.9%) for severe LRF impairment diagnosis (Table 5 and Figure 2).

DISCUSSION

So far, how to accurately evaluate LRF has been a hot topic in domestic and overseas research. As classic scoring systems, CTP score and MELD score had been widely used in clinical practice. The CTP introduced an element of bias into the scoring system due to the subjective nature of how clinical encephalopathy and ascites variables may be graded[19]. MELD score was a continuous variable, and each indicator was given corresponding weight through statistical analysis, which had further accuracy in evaluating LRF. In recent years, ALBI and PTAR models had been gradually applied in clinical practice, which were better to evaluate the LRF[20,21]. However, the ICG clearance test was currently considered the most valuable test in assessing LRF.

Although the ICG clearance test was a simple and helpful tool to assess individual LRF, it was an invasive and complex procedure, and its result was influenced by many factors (such as biliary excretion disorder and low proteinemia). In particular, the ICG clearance test was not applicable for pregnant women, the patients with a history of iodine allergy or hyperthyroidism[22]. Transient elastography (TE) was a non-invasive and reproducible technique for assessing liver fibrosis, even as a replacement for liver biopsy[23,24]. The Baveno VII Consensus showed that TE was an accurate tool for prediction of CSPH[25]. Meanwhile, in the previous study, it was found that LSM could
predict postoperative liver failure in patients with HCC\textsuperscript{26}. Therefore, LSM was considered to have a strong relationship with liver function.

Since liver function impairment was the primary determinant of post-hepatectomy liver failure development, the vast majority of candidates to Liver resection belonged to CTP A\textsuperscript{27}. According to the CTP classification, the majority of patients with HCC were classified as grade A, but their liver function may vary significantly\textsuperscript{15}. A previous study revealed that ICGR15 was more accurate than the CTP score and MELD score in predicting hepatic functional reserve before hepatectomy\textsuperscript{3}. The study showed that ICGR15 > 15% was a accurate method to predict postoperative hepatic decompensation in patients with CTP A\textsuperscript{28}. In patients with an ICGR15 > 20%, the previous study recommended nonanatomical resection rather than anatomical resection for the treatment of a solitary 2–5-cm-diameter HCC and without macroscopic vascular invasion\textsuperscript{29}. Therefore, it was essential to assess LRF before HCC hepatectomy, thereby assisting clinical decision-making.

Our research conducted new models for clinical prediction of LRF impairment based on LSM, and the models were superior to the other existing methods in predicting LRF (Table 4 and Figure 1). Moreover, compared to the other four methods, the models also showed better performance for predicting LRF in the prospective validation cohort (Table 5 and Figure 2). Therefore, based on the analysis of the above research results, the model could become an alternative tool for LRF assessment, especially in evaluating a population almost entirely stratified as CTP A.

\textit{Limits of the study}

Despite the significant findings in this study, our research also had a few limitations. First, the study was limited by its single-center prospective cohort nature. The patients were recruited from the same medical facility, and not all patients completed clinical data were obtained from a treatment database. Second, most of the patients in this study were Asians with B viral hepatitis. Therefore, the performance of the model in patients of other ethnicities (e.g., Caucasians, Africans, etc.) still needs further
investigation. Third, this model was mainly used to evaluate LRF in patients with compensated chronic liver disease, and its predictive value in the patient with decompensated stage need to be further evaluated. Finally, although the formula for the model was relatively complex, a mobile app or web-based calculator could calculate the score easily and rapidly in the current high-tech era. Despite these limitations, this study provided the first accurate model for China to evaluate LRF based on LSM.

CONCLUSION

The first predicted model based on LSM could facilitate accurate, reliable and simple-to-use prediction of the LRF irrespective of aetiology. It was entirely objective based on routine clinical and laboratory parameters. The model would be a useful tool for realizing individualized LRF evaluation to improve the popularity of testing and avoid possible risks during the ICG clearance test, ultimately achieving a clinically feasible and safe LRF test.

ARTICLE HIGHLIGHTS

Research background

There are no obvious clinical symptoms of the chronic liver disease (CLD) patients at an early stage, but their liver reserve function (LRF) may be impaired. Early evaluation of LRF is of great help for identifying disease progression Assessment of LRF is essential to predict prognosis for patients with CLD and determines the extent of liver resection patients with hepatocellular carcinoma (HCC).

Research motivation

Since liver function impairment was the primary determinant of post-hepatectomy liver failure development. There are no obvious clinical symptoms of the CLD patients at Child-Turcotte-Pugh A stage, but their LRF may be impaired. Due to the impossible implementation of the indocyanine green (ICG) clearance test in some CLD patients, we need to explore a new method to assess LRF accurately.
Research objectives
This study aimed to establish noninvasive models of LRF assessment based on LSM. The new predictive models were established through logistic regression analysis and were validated internally in a prospective cohort. The new models had a good predictive performance on LRF and could replace ICG clearance test, especially for the patients who were unable to undergo ICG testing.

Research methods
Clinical data from 360 patients with compensated CLD were retrospectively collected and analyzed in a training cohort. The new predictive models were established through logistic regression analysis and were validated internally in a prospective cohort (132 patients). The areas under the curve (AUCs) were measured and compared to evaluate the discrimination ability of different models.

Research results
Our study defined indocyanine green retention rate at 15 min (ICGR15) ≥ 10% as mildly impaired LRF and ICGR15 ≥ 20% as severely impaired LRF. We constructed predictive models of LRF, called the mLPaM and sLPaM, which involved only LSM, prothrombin time international normalized ratio to albumin ratio, age and model for end-stage liver disease. The AUC of mLPaM model (0.855, 0.872, respectively) and sLPaM model (0.869, 0.876, respectively) were higher than that of other methods in the two cohorts. Meanwhile, the new models showed good sensitivity and accuracy for LRF impairment diagnosis in the validation cohort.

Research conclusions
Our study found that the new models had a good predictive performance on LRF and could replace ICG clearance test, especially for the patients who were unable to undergo ICG testing.
Research perspectives

This was not a multicenter study and the most of CLD patients in this study were Asians. Therefore, a multi-center prospective cohort study would evaluate further the performance of the predictive models, and the models in patients of other ethnicities still need further investigation. Meanwhile, the predictive value of the models in the patient with a decompensated stage need to be further evaluated.

ACKNOWLEDGEMENTS

All patients participated in this study were appreciated.

REFERENCES


<table>
<thead>
<tr>
<th>#</th>
<th>Source Description</th>
<th>Words</th>
<th>% Similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>researchonline.gcu.ac.uk Internet</td>
<td>68</td>
<td>2%</td>
</tr>
<tr>
<td>2</td>
<td>Yan-Yan Wang, Xin-Hua Zhao, Liang Ma, Jia-Zhou Ye, Fei-Xiang Wu, Juan Tang, Xue-Mei You, Bang-De Xiang, Le-Qun Li. &quot;Comparison of the ability of Child-Pugh score, MELD score, and ICG-R15 to assess preoperative hepatic functional reserve in patients with hepatocellular carcinoma&quot;, Journal of Surgical Oncology, 2018</td>
<td>38</td>
<td>1%</td>
</tr>
<tr>
<td>3</td>
<td>&quot;UEG Week 2018 Poster Presentations&quot;, United European Gastroenterology Journal, 2018</td>
<td>32</td>
<td>1%</td>
</tr>
<tr>
<td>5</td>
<td><a href="http://www.frontiersin.org">www.frontiersin.org</a> Internet</td>
<td>27</td>
<td>1%</td>
</tr>
<tr>
<td>6</td>
<td>bsdwebstorage.blob.core.windows.net Internet</td>
<td>25</td>
<td>1%</td>
</tr>
<tr>
<td></td>
<td>Author(s)</td>
<td>Title</td>
<td>Publication</td>
</tr>
<tr>
<td>---</td>
<td>-----------</td>
<td>-------</td>
<td>-------------</td>
</tr>
<tr>
<td>7</td>
<td>Beverley Kok, Juan Abraldes.</td>
<td>&quot;Child–Pugh Classification: Time to Abandon?&quot;</td>
<td>Seminars in Liver Disease, 2019</td>
</tr>
<tr>
<td>8</td>
<td>f6publishing.blob.core.windows.net</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Koichiro Haruki, Hiroaki Shiba, Nobuhiro Saito, Takashi Horiuchi et al.</td>
<td>&quot;Risk stratification using a novel liver functional reserve score of combination prothrombin time–international normalized ratio to albumin ratio and albumin in patients with hepatocellular carcinoma&quot;</td>
<td>Surgery, 2018</td>
</tr>
<tr>
<td>10</td>
<td><a href="http://www.semanticscholar.org">www.semanticscholar.org</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td><a href="http://www.hindawi.com">www.hindawi.com</a></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Rong-yun Mai, Jie Zeng, Hua-ze Lu, Rong Liang et al.</td>
<td>&quot;Combining Aspartate Aminotransferase-to-Platelet Ratio Index with Future Liver Remnant to Assess Preoperative Hepatic Functional Reserve in Patients with Hepatocellular Carcinoma&quot;</td>
<td>Journal of Gastrointestinal Surgery, 2020</td>
</tr>
<tr>
<td>13</td>
<td>link.springer.com</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>