Development of non-alcoholic fatty liver disease scoring system among adult medical check-up patients: a large cross-sectional and prospective validation study

Cosmas Rinaldi A Lesmana¹,²
Levina S Pakasi¹
Sri Inggrini³
Maria L Aidawati³
Laurentius A Lesmana¹
¹Digestive Disease and GI Oncology Centre, Medistra Hospital, ²Department of Internal Medicine, Hepatobiliary Division, Cipto Mangunkusumo Hospital, University of Indonesia, ³Radiology Department, Medistra Hospital, Jakarta, Indonesia

Background: Non-alcoholic fatty liver disease (NAFLD) is the most common liver disease in the community. However, NAFLD remains undiagnosed in most people with limited access to imaging facilities in most developing countries.

Objective: To examine the prevalence of NAFLD and to develop the risk scoring model for predicting the presence of NAFLD among adult medical check-up patients.

Method: A large prospective cross-sectional study was conducted among medical check-up patients who underwent transabdominal ultrasound examination between January and December 2013 in Medistra Hospital, Jakarta. Data were obtained from the patients’ medical records. Logistic regression analyses were undertaken to identify the best combination of risk factors for predicting fatty liver using the backward (likelihood ratio) approach. The adjusted odds ratio and 95% confidence interval were estimated using the logistic regression coefficient. The prediction model was assessed using the receiver operating characteristic curve and the Hosmer–Lemeshow goodness-of-fit test and was validated on a new, prospective cohort. Statistical analysis was done using SPSS version 17.

Results: A total of 1,054 cases was included in this study. Fatty liver was present in 538 (51.0%) patients. Bivariate analyses found associations among fatty liver and several risk factors. Six risk factors were incorporated to build the final prediction model. All scores were summed up to obtain the total score. A probability equation was developed by applying linear regression analysis on the total score. The prediction model had good diagnostic performance with an area under the receiver operating characteristic curve = 0.833 (95% confidence interval = 0.809–0.857). The Hosmer–Lemeshow goodness-of-fit P-value was 0.232, which indicated the appropriateness of the logistic regression model to predict fatty liver. On the validation set, the scoring system proved to be moderately accurate and can potentially be applied to larger population setting.

Conclusion: The presence of fatty liver in NAFLD patients can be predicted using our proposed fatty liver scoring system.

Keywords: fatty liver, scoring model, ultrasound, community, developing countries, diagnostic performance

Background
Non-alcoholic fatty liver disease (NAFLD) is a spectrum of liver disease that may progress to liver cirrhosis and hepatocellular carcinoma. In Asia Pacific the prevalence of NAFLD ranges from 5%–30%, and Indonesia is one of the high prevalence countries of NAFLD.¹² The risk factors of NAFLD in Asian population might be similar with...
the Western countries because of the food habit and lifestyle changes over the recent years. Theoretically, NAFLD is associated with insulin resistance condition and can represent the hepatic manifestation of the metabolic syndrome (MS) which comprises obesity, dyslipidemia, hypertension, and glucose intolerance.4,5

Most NAFLD patients are asymptomatic and might be diagnosed incidentally during imaging study. The use of liver parameters such as AST or ALT is not always reliable. However, NAFLD could also be suspected in patients with the presence of MS. Studies in the United States revealed good associations between metabolic parameters and the presence of NAFLD.6,7

Transabdominal ultrasonography is a simple, widely available, and acceptable tool for first-line screening for the presence of fatty liver in clinical practice. It has high diagnostic accuracy for the diagnosis of fatty liver.8,9 However, transabdominal ultrasound is not routinely done for medical check-up (MCU) in most primary health care centers or peripheral hospitals. Therefore, it is tempting to know whether the presence of fatty liver can be predicted with all metabolic risk factors.

The aim of this study was to know the prevalence of fatty liver among MCU patients and to develop a scoring system to predict the presence of fatty liver.

Method

Study design and subjects

The study design was a prospective analytical cross-sectional study among MCU patients in Medistra Hospital between January and December 2013. These subjects mostly include employees of private companies in Jakarta or middle-class income people who would like a comprehensive MCU, which includes transabdominal ultrasound and blood chemistry testing for major cardiovascular risk factors. Medistra Hospital is one of the top private hospitals in Jakarta which shares the same patient economic profiles with other private hospitals.

The minimum sample size for estimating one population proportion at an anticipated proportion of 50%, 99% level of confidence, and 5% confidence interval (CI) was 664 subjects.

The study protocol was approved by the Internal Review Board of the hospital. Inclusion criteria were adult patient aged more than 18 years, having routine laboratory check-up including liver function test, fasting blood glucose (FBG) levels, lipid profile, alcohol consumption less than 20 g/day, and underwent transabdominal ultrasound assessment. Patients were excluded from analyses if laboratory data were incomplete based on medical records. Predictors of NAFLD tested were sex, age, smoking history, history of hypertension, history of diabetes, body mass index (BMI), systolic and diastolic blood pressures, FBG, triglycerides, total cholesterol, low-density lipoprotein-cholesterol, high-density lipoprotein (HDL)-cholesterol, serum AST, and ALT levels.

Diagnosis of fatty liver

Fatty liver was diagnosed by transabdominal ultrasound. Ultrasound assessment was performed by two senior radiologists in Medistra Hospital. Ultrasound was done using a 3.5 MHz convex-array transducer (LOGIC S6 and LOGIC 7, GE Health Care). Both radiologists who performed the ultrasound evaluation were blinded to the laboratory data. The diagnostic performance of ultrasound evaluation of fatty liver is high but not perfect.10-14 Therefore, the presence of fatty liver was confirmed by the ultrasound and liver laboratory results.

For statistical analyses, the ultrasound appearance of the liver was grouped into “fatty liver” and “no fatty liver”.

Data management and statistical analyses

Data were obtained by retrieving the medical records of the full 1-year entries of MCU patients. These data include comprehensive interview of the patients’ medical history, physical examination, laboratory data, and abdominal ultrasound results.

Logistic regression analyses were undertaken to identify the best combination of risk factors for predicting fatty liver using the backward (likelihood ratio) approach. The adjusted odds ratio and 95% CI were estimated using the logistic regression coefficient. The diagnostic performance of the prediction model was evaluated using the receiver operating characteristic curve with 95% CI, whereas the goodness-of-fit of the model was assessed by the Hosmer–Lemeshow statistics and pseudo-$R$-squares (Cox and Snell $R$-square and Nagelkerke $R$-square). Statistical analysis was done using SPSS version 17.0.

Validation study

The new fatty liver prediction model was also tested using a prospective data set from MCU patients between July and
December 2014. Clinical and laboratory data were recorded and scored accordingly. Diagnostic performance test was applied on the total score results.

Result

Characteristics of the study subjects

A total of 1,057 patients were recorded during the study period; three of them did not provide complete laboratory data leaving a final number of 1,054 cases for statistical analyses. Fatty liver was present in 538 (51.0%) patients, most of them were mild. There were 13 (1.2%) who already had a chronic parenchymal liver disease based on ultrasound imaging (Table 1).

Bivariate and multivariate analyses

Bivariate analyses found associations among NAFLD and several predictors, ie, male sex, age ≥35 years, BMI ≥25 kg/m², FBG >100 mg/dL, systolic blood pressure ≥130 mmHg, diastolic blood pressure ≥85 mmHg, triglyceride levels ≥150 mg/dL, total cholesterol ≥200 mg/dL; HDL-cholesterol ≤40 (for men) or ≤50 mg/dL (for women), and serum aminotransferase levels ≥35 U/L. All variables showed significant association with the presence of NAFLD (Table 2) and were entered into subsequent multivariate analysis. The resulting independent predictors for NAFLD were male sex, age ≥35 years, BMI ≥25 kg/m², triglyceride levels ≥150 mg/dL; HDL-cholesterol <40 mg/dL (for men) or <50 mg/dL (for women), and serum ALT levels ≥35 U/L (Table 3).

Development of fatty liver index

Variables for fatty liver index development were chosen from independent predictors which have adjusted odds ratio more than 1.5 on multivariate analyses. Scoring was calculated by first dividing each of β value to its standard error which resulted in standardized β (β/standard error) values. The lowest value of standardized β was then marked and divided by itself to yield a score of 1.0. Other standardized β values were then divided to that lowest values and resulted as score for each variable. Each score was rounded to one decimal digit (Table 3).

There were six predictors included in the scoring system, ie, sex, age, BMI, triglyceride level, HDL-cholesterol level, and serum ALT level (Table 3). To simplify the probability calculation, all scores were summed up to yield a total score. The linear regression equation was:

$$y = -2.568 + (0.478 \times \text{Total score})$$

(Table 4), and probability of having fatty liver can be calculated as follows:

$$Pr = \frac{1}{1 + e^{(-2.568 + 0.478 \times \text{Total score})}}$$

Diagnostic accuracy of the equation showed that the area under the receiver operating characteristic curve of the prediction model was 0.833 (95% CI = 0.809–0.857; P<0.001) (Figure 1). The Hosmer–Lemeshow goodness-of-fit P-value was 0.232, which indicated the appropriateness of the logistic regression model to predict fatty liver. However, the extent of this model to predict fatty liver in the study population was moderate (Cox and Snell R-square = 0.317; Nagelkerke R-square = 0.422). Diagnostic performance of the prediction model was good (Table 5).

Validation study

There were 454 patients (258 men and 196 women) enrolled in this validation phase. Fatty liver was diagnosed in 208 (45.8%) of them. By using our prediction model, the AUC to detect fatty liver was 0.822 (95% CI: 0.784–0.860; P<0.001). Diagnostic performance was comparable with the prediction set with moderate sensitivity and specificity (Table 5).

Discussion

To our knowledge, this is the first study in Indonesia to develop a scoring system to predict NAFLD in adults. Large sample size on the prediction set is a strength of this study which was further validated on a new, prospective cohort. Subjects enrolled in this study were mostly healthy.
Hypertriglyceridemia and low HDL-cholesterol levels are also commonly found in patients with NAFLD with an estimated prevalence of 50%.

It has been demonstrated by several studies that higher risk of metabolic disorders is found in patients with NAFLD than in those without it. However, since data on metabolic disorders and NAFLD were collected and analyzed cross-sectionally, the causal association between metabolic disorders and NAFLD is unclear. Moreover, although there was a significant association between diabetes and NAFLD, neither a history of diabetes nor high FBG served as an independent risk factor for NAFLD on multivariate analyses suggesting that diabetes is a confounding factor. This could imply that diabetes share the same risk profile with fatty liver but not necessarily be a risk factor if present alone.

Elevated ALT activity levels is commonly associated with NASH, although normal ALT does not rule out significant liver disease in NAFLD patients. Furthermore, the upper normal limit (ULN) of ALT has been recently evaluated. It has been suggested to lower the ULN, especially in women, but regardless of the cutoff, investigators found that the prevalence of NAFLD and NASH in normal ALT patients

<table>
<thead>
<tr>
<th>Risk factor</th>
<th>NAFLD</th>
<th>No NAFLD</th>
<th>OR (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male sex (n=720)</td>
<td>444 (61.7%)</td>
<td>276 (38.3%)</td>
<td>4.107 (3.098–5.445)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age &gt;35 years (n=723)</td>
<td>425 (58.8%)</td>
<td>298 (41.2%)</td>
<td>2.751 (2.098–3.609)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Smoking history (n=242)</td>
<td>138 (57.0%)</td>
<td>115 (24.0%)</td>
<td>1.360 (1.018–1.817)</td>
<td>0.037</td>
</tr>
<tr>
<td>Hypertension (n=140)</td>
<td>103 (73.6%)</td>
<td>37 (26.4%)</td>
<td>3.065 (2.060–4.561)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Systolic BP ≥130 mmHg (n=251)</td>
<td>169 (67.3%)</td>
<td>82 (32.7%)</td>
<td>2.424 (1.799–3.266)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Diastolic BP ≥85 mmHg (n=139)</td>
<td>101 (72.2%)</td>
<td>38 (27.3%)</td>
<td>2.907 (1.959–4.315)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>BMI ≥25 kg/m² (n=493)</td>
<td>373 (75.7%)</td>
<td>120 (24.3%)</td>
<td>7.460 (5.670–9.815)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Diabetes (n=80)</td>
<td>59 (73.8%)</td>
<td>21 (26.2%)</td>
<td>2.903 (1.737–4.853)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>FBG ≥100 mg/dL (n=227)</td>
<td>159 (70.0%)</td>
<td>68 (30.0%)</td>
<td>2.764 (2.017–3.788)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Triglycerides ≥150 mg/dL (n=280)</td>
<td>210 (75.0%)</td>
<td>70 (25.0%)</td>
<td>4.079 (3.005–5.538)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Total cholesterol ≥200 mg/dL (n=564)</td>
<td>328 (58.2%)</td>
<td>236 (41.8%)</td>
<td>1.853 (1.451–2.367)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>HDL &lt;40 (M) or &lt;50 (F) mg/dL (n=237)</td>
<td>159 (67.1%)</td>
<td>78 (32.9%)</td>
<td>2.356 (1.739–3.192)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>LDL-cholesterol ≥100 mg/dL (n=873)</td>
<td>469 (53.7%)</td>
<td>404 (46.3%)</td>
<td>1.821 (1.357–2.366)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Serum AST level ≥35 U/L (n=81)</td>
<td>62 (76.5%)</td>
<td>19 (23.5%)</td>
<td>3.400 (2.003–5.772)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Serum ALT level ≥35 U/L (n=241)</td>
<td>191 (79.3%)</td>
<td>50 (20.7%)</td>
<td>5.119 (3.640–7.199)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Abbreviations: LDL, low-density lipoprotein; HDL, high-density lipoprotein; NAFLD, non-alcoholic fatty liver disease; BMI, body mass index; TG, triglycerides; OR, odds ratio; CI, confidence interval; FBG, fasting blood glucose; BP, blood pressure; M, male; F, female; AST, aspartate transaminase; ALT, alanine transaminase.

Table 3 Independent risk factors to predict fatty liver (N=1,054)

<table>
<thead>
<tr>
<th>Risk factor</th>
<th>β</th>
<th>SE (β)</th>
<th>β/SE</th>
<th>OR (95% CI)</th>
<th>P-value</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male sex</td>
<td>0.943</td>
<td>0.169</td>
<td>5.567</td>
<td>2.567</td>
<td>1.842–3.577</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age &gt;35 years</td>
<td>1.049</td>
<td>0.172</td>
<td>6.089</td>
<td>2.855</td>
<td>2.037–4.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>BMI ≥25 kg/m²</td>
<td>1.565</td>
<td>0.153</td>
<td>10.207</td>
<td>4.781</td>
<td>3.540–6.457</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Triglycerides ≥150 mg/dL</td>
<td>0.819</td>
<td>0.184</td>
<td>4.450</td>
<td>2.268</td>
<td>1.581–3.254</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>HDL &lt;40 (M) or &lt;50 (F) mg/dL</td>
<td>0.546</td>
<td>0.191</td>
<td>2.855</td>
<td>1.726</td>
<td>1.187–2.511</td>
<td>0.004</td>
</tr>
<tr>
<td>ALT ≥35 U/L</td>
<td>1.169</td>
<td>0.202</td>
<td>5.776</td>
<td>3.220</td>
<td>2.165–4.788</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.625</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: BMI, body mass index; HDL, high-density lipoprotein; M, male; F, female; SE, standard error; OR, adjusted odds ratio; CI, confidence interval; ALT, alanine transaminase.
Table 4 Linear regression model of the total score to predict fatty liver

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coefficient (β)</th>
<th>Standard Error (SE(β))</th>
<th>Adjusted OR (ORadj)</th>
<th>95% CI</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total score</td>
<td>0.478</td>
<td>0.029</td>
<td>1.612</td>
<td>1.522–1.708</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Constant</td>
<td>−2.568</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** SE, standard error; ORadj, adjusted odds ratio; CI, confidence interval.

Table 5 Diagnostic performance of total score for fatty liver at a cutoff of 5.0

<table>
<thead>
<tr>
<th>Parameter</th>
<th>On prediction set (95% CI)</th>
<th>On validation set (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnostic parameter</td>
<td>(n=1,054)</td>
<td>(n=454)</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>79.0% (75.3–82.4%)</td>
<td>75.5% (69.1–81.2%)</td>
</tr>
<tr>
<td>Specificity</td>
<td>72.2% (68.1–76.1%)</td>
<td>70.3% (64.2–76.0%)</td>
</tr>
<tr>
<td>PPV</td>
<td>74.8% (71.0–78.3%)</td>
<td>68.3% (61.8–74.2%)</td>
</tr>
<tr>
<td>NPV</td>
<td>76.7% (72.7–80.4%)</td>
<td>77.2% (71.2–82.6%)</td>
</tr>
<tr>
<td>Positive LR</td>
<td>2.84 (2.46–3.29)</td>
<td>2.54 (2.07–3.13)</td>
</tr>
<tr>
<td>Negative LR</td>
<td>0.29 (0.24–0.35)</td>
<td>0.35 (0.27–0.45)</td>
</tr>
<tr>
<td>Disease prevalence</td>
<td>51.1% (48.0–54.2%)</td>
<td>45.8% (41.2–50.5%)</td>
</tr>
<tr>
<td>Kappa statistics</td>
<td>0.519 (P&lt;0.001)</td>
<td>0.454 (P&lt;0.001)</td>
</tr>
</tbody>
</table>

**Abbreviations:** CI, confidence interval; PPV, positive predictive value; NPV, negative predictive value; LR, likelihood ratio.

was similar. For simplicity, we do not distinguish the ULN for men and women in our study.

When applying our prediction model to the new validation set, there is slight reduction in diagnostic performance of the scoring system. This could be due to the small number of subjects enrolled and slightly lower disease prevalence in the validation set. Overall, the scoring system is moderately accurate in detecting NAFLD and is potentially applicable in large, population-based screening. Screening of fatty liver in adult population visiting primary care clinics has not been advised due to lack of clinical importance and limited treatment choices. In addition, screening with liver biochemistry tests may not be sufficiently sensitive, whereas liver ultrasound imaging may not be practical. However, with our scoring system, any person who underwent routine basic cardiovascular check-up could benefit from knowing his/her chance of having fatty liver without the need to ask for imaging studies. The scoring system produced by this study include simple clinical factors that could be assessed anywhere in the primary health care institution, ie, age, sex, BMI, triglycerides, HDL-cholesterol, and ALT levels. By using this scoring system, most clinicians or any health care providers will be able to do the screening in a large population. Although benefit of screening may not be clearly seen at the moment due to uncertainties of the true liver condition (whether it is a simple steatosis or a NASH); in our opinion, individuals with higher fatty liver scores calculated by our scoring system should be advised regarding more healthy living habits such as calorie restriction and exercise.

There are limitations that should be considered based on this study. Firstly, diagnosis of NAFLD was made by ultrasound. Considering the technique used for liver imaging, computed tomography and magnetic resonance imaging are actually more accurate to detect NAFLD than ultrasound. However, both computed tomography and magnetic resonance imaging are more expensive, more time consuming, and cannot be used as part of routine MCU. Secondly, there was no histopathological assessment to confirm the diagnosis of fatty liver, as ultrasound imaging can only detect steatosis if there is more than 33% fat content in the liver. Consequently, fatty liver may still be present in patients with negative results on ultrasound imaging. However, ultrasound is still a recommended technique for use in a large population for screening of fatty liver because it is relatively low-cost, non-invasive, and widely available. A meta-analysis has supported the use of ultrasound for NALFD diagnosis because it has good sensitivity (84.8%) and specificity (93.6%). In addition, liver biopsy is an invasive procedure which cannot be done on routine MCU or screening programs.

**Conclusion**

The prevalence of NAFLD among MCU patients in our hospital is high (51%). The screening program using our simple model score might be very useful in daily practice.

![Figure 1 ROC curve of predicted total score for fatty liver](AUC = 0.833; 95% CI: 0.809–0.857; P<0.001).

**Abbreviations:** ROC, receiver operating characteristic; AUC, area under the curve; CI, confidence interval.
especially in primary health care centers. On the validation set, the scoring system was proved to be moderately accurate and can potentially be applied to larger population setting.

Compliance with ethical requirements

All procedures followed in this study were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2008.

Disclosure

The authors have no conflicts of interest to disclose.

References