The albumin–bilirubin score as a predictor of outcomes in Japanese patients with PBC: an analysis using time-dependent ROC

Takanori Ito1*, Masatoshi Ishigami2, Hikaru Morooka3, Kenta Yamamoto1, Norihiro Imai1, Yoji Ishizu1, Takashi Honda1, Daisaku Nishimura2, Toshifumi Tada1, Satoshi Yasuda4, Hidenori Toyoda4, Takashi Kumada5 & Mitsuhiro Fujishiro1

The albumin–bilirubin (ALBI) score is calculated using only serum albumin and bilirubin levels, and was developed as a simple method to assess hepatic function. In this study, a total of 409 patients with primary biliary cholangitis (PBC) were enrolled between March 1990 and October 2018. The predictive performances of the ALBI score and other well-established prognostic scores were compared using time-dependent receiver operating characteristic (ROC) analysis. During the follow-up period, 60 patients died, 45 due to liver-related diseases and 15 due to non-liver-related diseases, and 16 patients underwent liver transplantation. Time-dependent ROC analysis showed that the ALBI score has higher the areas under the ROC curves (AUROCs) than the Child–Pugh (C–P) score at each time point; AUROCs at 3, 5, and 10 years after the start of follow-up were 0.94, 0.91, and 0.90 for the ALBI score, and 0.89, 0.88, and 0.82 for the C–P score, respectively. The ALBI score showed the highest AUROCs within 2 years after the start of observation; beyond 2 years, however, the Mayo score had better prognostic ability for mortality and liver transplantation. The ALBI score/grade, derived from objective blood tests, and the Mayo score were superior prognostic tools in PBC patients.

Abbreviations
ALBI  Albumin–bilirubin
PBC  Primary biliary cholangitis
ROC  Receiver operating characteristic
AUROC  Area under the receiver operating characteristic curve
C–P  Child–Pugh
MELD  Model of end-stage liver disease
AASLD  American Association for the Study of Liver Diseases
AIH  Autoimmune hepatitis
HCC  Hepatocellular carcinoma
ULN  Upper limit of normal
AST  Aspartate aminotransferase
ALT  Alanine aminotransferase
γ-GTP  γ-Glutamyl transpeptidase
ALP  Alkaline phosphatase
SD  Standard deviation
AIC  Akaike’s information criterion
HR  Hazard ratio

1Department of Gastroenterology and Hepatology, Nagoya University Graduate School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya 466-8550, Japan. 2Department of Gastroenterology, Toyota Kosei Hospital, Toyota, Japan. 3Department of Internal Medicine, Himeji Red Cross Hospital, Himeji, Japan. 4Department of Gastroenterology and Hepatology, Ogaki Municipal Hospital, Gifu, Japan. 5Department of Nursing, Gifu Kyoritsu University, Gifu, Japan. *email: tahkun56@gmail.com
Primary biliary cholangitis (PBC) is an autoimmune chronic liver disease, which is characterized by destruction of the intrahepatic bile ducts, resulting in cholestasis and progressive fibrosis. The majority of PBC patients are asymptomatic at the time of diagnosis, and the clinical course is typically slow and progressive, but can differ significantly among individuals. Although the overall 5-year survival rate of patients with PBC is 80–90%, a significant proportion of patients suffer from cirrhosis-related complications or hepatic malignancies.

To distinguish patients at high risk of developing complications, several non-invasive prognostic markers have been used in patients with PBC, such as the Child–Pugh (C–P) score/classification, the model of end-stage liver disease (MELD), the Mayo risk score, and the Newcastle model. All of these markers include the serum bilirubin level; in fact, the guidelines of the American Association for the Study of Liver Diseases (AASLD) state that serum bilirubin level is the most important factor for predicting prognosis in patients with PBC.

The C–P classification system is one of the most commonly used methods worldwide for assessing hepatic function. However, the C–P score/classification system includes subjective components, such as ascites and encephalopathy, and interrelated factors such as serum albumin and ascites. Recently, a simple new method was developed to assess hepatic function; known as the albumin–bilirubin (ALBI) grade, it is calculated using only serum albumin and bilirubin values.

However, it is not yet clear whether the ALBI score/grade can serve as a novel biomarker capable of predicting prognosis in patients with PBC. Therefore, in this study, we evaluated the impact of several serum prognostic markers, including the ALBI score/grade, for predicting the prognosis of Japanese patients with PBC in the short and long term, using time-dependent receiver operating characteristic (ROC) analysis.

**Results**

**Patient characteristics.** The characteristics of the study patients are shown in Table 1. The mean age was 59.5 years, and there was a predominance of females (83.4%). The mean values of the ALBI score, C–P score, Mayo risk score, MELD score, and FIB-4 index were −2.59, 5.9, 5.10, 8.06, and 2.26, respectively. The mean follow-up period was 8.8 years. Almost all patients (n = 379, 92.7%) were treated with ursodeoxycholic acid (UDCA; 600–900 mg per day), and 41 patients (10.0%) who had already undergone UDCA treatment received corticosteroids. These five patients were treated with corticosteroids due to the comorbidities (rheumatoid arthritis, n = 3; Sjögren's syndrome, n = 1; and aplastic anemia, n = 1). One hundred eighty-five (45.2%) had the history of liver biopsy, and Scheuer's stage was I for 138 cases (74.6%), II for 31 cases (16.8%), III for 14 cases (7.5%), and IV for 2 cases (1.1%). Only 19 patients (4.6%) with a history of HCC met the Milan Criteria or experienced a complete cure. During the follow-up period, 60 (14.7%) patients died, 45 due to liver-related disease and 15 due to non-liver-related disease, and 16 (3.9%) patients underwent liver transplantation.

**Cumulative survival rate, incidence rates of death and liver transplantation, and causes of death.** The overall and liver transplantation-free survival rates of patients at 5, 10, 15, and 20 years were 87.8% (95% confidence interval [CI], 83.9–90.8%), 80.0% (95% CI, 74.9–84.2%), 73.9% (95% CI, 67.6–79.2%), and 70.5% (95% CI, 63.2–76.6%), respectively (Supplementary Fig. S1 online). The 1-, 5-, 10-, 15-, and 20-year cumulative incidence rates of all-cause death were 2.3% (95% CI, 1.1–4.1%), 8.3% (95% CI, 5.7–11.5%), 16.1 (95% CI, 12.0–20.7%), 21.7 (95% CI, 16.4–27.4%), and 25.0% (95% CI, 18.8–31.8%); those of liver transplantation were 2.8% (95% CI, 1.5–4.7%), 3.9% (95% CI, 2.3–6.2%), 3.9% (95% CI, 2.3–6.2%), 4.4% (95% CI, 2.6–7.0%), and 4.4% (95% CI, 2.6–7.0%) (Fig. 1a). The 5-, 10-, 15-, and 20-year cumulative incidence rates of liver-related death were 6.2% (95% CI, 3.9–9.1%), 13.8% (95% CI, 9.9–18.3%), 16.0% (95% CI, 11.5–21.1%), and 19.5% (95% CI, 13.8–26.0%); those of non-liver-related death were 2.4% (1.1–4.5%), 2.9% (95% CI, 1.4–5.2%), 6.5% (95% CI, 3.5–10.9%), and 6.5% (95% CI, 3.5–10.9%) (Fig. 1b). Of patients with liver-related deaths, 13 died due to HCC and 32 died due to liver cirrhosis-related complications other than HCC. In the remaining 71 patients with non-liver-related deaths, nine died due to extra-hepatic malignancies and eight died of other causes (Supplementary Table S2 online).

**Factors associated with prognosis by Fine-Grey proportional hazards models.** Table 2 shows that the following parameters were selected as significant risk factors: age, C–P score, ALBI score, FIB-4 index, MELD score, and Mayo risk score for liver-related death and liver transplantation; age, male, C–P score, ALBI score, FIB-4 index, and Mayo risk score for non-liver-related death. Univariate analysis with Fine-Grey proportional hazards models showed most of prognostic factors used in this study were selected as independent risk factors associated both liver-/non-liver-related deaths in PBC.

**Time-dependent ROC analysis for overall survival and the incidence of liver transplantation.** Figure 2 shows the plots of AUROCs for the ALBI score, C–P score, FIB-4 index, Mayo score, and MELD score for patient overall and liver transplantation-free survival from 1 to 10 years after the start of follow-up. The detailed ROC curves for each marker after the start of follow-up, obtained using time-dependent ROC analysis, are shown in Supplementary Fig. S3 online. Time-dependent ROC analysis showed that the predictive power of the ALBI score for overall mortality and the incidence of liver transplantation was superior to that of the C–P score for all years; the AUROCs at 3, 5, 7, and 10 years after the start of follow-up were 0.941, 0.906,
Table 1. Patient characteristics. Continuous variables are expressed as mean ± SD (standard deviation). AST aspartate aminotransferase; ALT alanine aminotransferase; γ-GTP γ-glutamyl transpeptidase; ALP alkaline phosphatase; ALBI albumin–bilirubin; FIB-4 Fibrosis-4; MELD model of end-stage liver disease; UDCA ursodeoxycholic acid.

|                                | (n = 409)       |
|--------------------------------|-----------------|
| Age (years)                    | 59.5 ± 12.7     |
| Sex (female/male)              | 341 (83.4%)/68 (16.6%) |
| Platelet count (× 10⁶/m³)      | 21.3 ± 10.03    |
| AST (U/L)                      | 79 ± 166        |
| ALT (U/L)                      | 87 ± 223        |
| γ-GTP (U/L)                    | 87 ± 223        |
| ALP (U/L)                      | 609 ± 446       |
| Total bilirubin (mg/dL)        | 1.39 ± 2.87     |
| Albumin (g/dL)                 | 3.92 ± 0.59     |
| Prothrombin time (%)           | 99.0 ± 22.7     |
| Immunoglobulin M (mg/dL)       | 436 ± 366       |
| Anti-mitochondrial antibody positive (%) | 373 (91.2%)     |
| Child–Pugh classification (A/B/C) | 321 (78.5%)/62 (15.2%)/26 (6.4%) |
| Child–Pugh score               | 5.9 ± 1.7       |
| ALBI score                     | −2.59 ± 0.65    |
| FIB-4 index                    | 2.26 ± 4.05     |
| MELD score                     | 8.06 ± 3.44     |
| Mayo risk score                | 5.10 ± 2.00     |
| UDCA/bezafibrate/steroid use   | 379 (92.7%)/122 (29.8%)/5 (1.2%) |
| History of liver biopsy (+ / −) | 185 (45.2%)/224 (54.8%) |
| Scheuer's stage (I/II/III/IV)   | 138 (74.6%)/31 (16.8%)/14 (7.57%)/2 (1.1%) |
| History of hepatocellular carcinoma (+ / −) | 19 (4.6%)/390 (95.4%) |
| Event (death/liver transplantation) | 60 (14.7%)/16 (3.9%) |
| Liver-related/non-liver-related deaths | 45 (11.0%)/15 (3.7%) |
| Observation period (years)      | 8.8 ± 7.1       |

Figure 1. Incidence rates of liver-related/non-liver-related death and liver transplantation in all patients. (a) Cumulative incidence rates were 2.3%, 8.3%, 16.1%, 21.7%, and 25.0% for all-cause death, and 2.8%, 3.9%, 3.9%, 4.4%, and 4.4% for liver transplantation at 1, 5, 10, 15, and 20 years, respectively. (b) Cumulative incidence rates were 6.2%, 13.8%, 16.0%, and 19.5% for liver-related death, and 2.4%, 2.9%, 6.5%, and 6.5% for non-liver-related death at 5, 10, 15, and 20 years, respectively.
0.916, and 0.896 for the ALBI score, and 0.893, 0.876, 0.854, and 0.82 for the C–P score, respectively. Over the course of the follow-up period, the Mayo risk score gradually became superior to the other markers, with the highest AUROCs for mortality and liver transplantation; however, the value of AUROCs in ALBI score showed larger than the Mayo risk score within 2 years after the start of observation. The AUROCs of the FIB-4 index were the lowest of all the markers.

**Table 2.** Factors associated with prognosis with Fine-Grey proportional hazards model. HR hazard ratio; CI confidence interval; ALBI albumin–bilirubin; FIB-4 Fibrosis-4; MELD model of end-stage liver disease.

| Variables          | Liver-related death or liver-transplantation | Non liver-related death |
|--------------------|---------------------------------------------|-------------------------|
|                    | HR (95% CI)                                 | P value                 | HR (95% CI)                                 | P value                 |
| Age                | 1.065 (1.041–1.090)                          | <0.001                  | 1.124 (1.056–1.197)                          | <0.001                  |
| Gender; female     | 1                                           |                         | 1                                           |                         |
| Male               | 1.019 (0.679–1.529)                          | 0.930                   | 1.940 (1.159–3.248)                          | 0.012                   |
| Child–Pugh score   | 1.798 (1.580–2.047)                          | <0.001                  | 1.243 (1.014–1.524)                          | 0.036                   |
| ALBI score         | 5.521 (3.858–7.900)                          | <0.001                  | 2.331 (1.387–3.917)                          | 0.001                   |
| FIB-4 index        | 1.046 (1.015–1.078)                          | 0.003                   | 1.047 (1.014–1.081)                          | 0.005                   |
| MELD score         | 1.199 (1.118–1.285)                          | <0.001                  | 1.083 (0.995–1.173)                          | 0.051                   |
| Mayo risk score    | 1.748 (1.531–1.997)                          | <0.001                  | 1.350 (1.168–1.559)                          | <0.001                  |

**Figure 2.** Time-dependent AUROCs for overall and liver transplantation-free survival after the start of follow-up. The albumin–bilirubin (ALBI) score demonstrated higher area under the receiver operating characteristic curves (AUROCs) for survival and liver transplantation than the Child–Pugh score at each time point. The Mayo risk score kept consistently higher AUROCs for outcome during the follow-up period; however, the ALBI score showed higher AUROCs than the Mayo risk score within 2 years after the start of observation. The AUROC of the Fibrosis-4 index was the lowest of all the markers. AUROC area under the receiver operating characteristic curve; ALBI albumin–bilirubin; FIB-4 Fibrosis-4; MELD model of end-stage liver disease.

Overall survival and liver transplantation-free rates based on the ALBI grade and C–P class. In patients classified as ALBI-grade 1 (n = 274), 2 (n = 105), and 3 (n = 30), the survival and liver transplantation-free rates were 97.7% (95% CI, 94.4–99.0%), 80.0% (95% CI 70.0–87.0%), and 24.3% (95% CI, 10.2–41.6%) at...
5 years, 95.1% (95% CI, 90.6–97.4%), 59.3% (95% CI, 45.8–70.5%), and 6.5% (95% CI, 0.5–24.2%) at 10 years, and 87.5% (95% CI, 80.1–92.2%), 55.6% (95% CI, 41.0–67.9%), and 6.5% (95% CI, 0.5–24.2%) at 15 years, respectively (Fig. 3a). In addition, in patients categorized as C–P class A (n = 321), B (n = 62), and C (n = 26), the survival and liver transplantation-free rates were 97.3% (95% CI, 94.4–98.7%), 65.0% (95% CI, 51.0–75.9%), and 27.5% (95% CI, 11.4–46.4%) at 5 years, 91.2% (95% CI, 86.2–94.4%), 51.7% (95% CI, 36.5–65.0%), and 14.6% (95% CI, 3.1–34.5%) at 10 years, and 85.9% (95% CI, 79.1–90.6%), 39.8% (95% CI, 23.8–55.4%), and 14.6% (95% CI, 3.1–34.5%) at 15 years, respectively (Fig. 3b). Both the ALBI grade and C–P classification showed good discriminatory ability for prognosis (mortality and liver transplantation). However, the AIC of the ALBI grade was better than that of the C–P classification (708.96 vs. 720.37).

Discussion

Although several prognostic scores have been reported to be associated with prognosis in PBC patients, it was unclear whether these markers reflect prognosis in both the short and long term. Therefore, we compared these well-established predictive markers, including the ALBI score, using time-dependent ROC curves in PBC patients. In our cohort, although more than 70% of deaths in patients with PBC were due to liver-related disease, time-dependent ROC analysis showed that the ALBI score and Mayo score had higher AUROCs than the other markers for predicting overall survival and the incidence of liver transplantation.

ROC analysis is commonly used to evaluate the discriminatory power of a continuous variable for a binary disease outcome. However, it is impossible to compare the prognosis determined using general ROC analysis because outcomes are time dependent. In addition, the Kaplan–Meier method, which is frequently used to investigate prognosis, requires optimal cutoff values for each marker. For instance, there is no established cutoff value for the Mayo score, and it is unclear whether the range of FIB-4 index cutoff values from 1.45 to 3.25, as determined in a previous study on hepatitis C virus infection, is applicable to PBC. On the other hand, time-dependent ROC curves have been introduced for assessing the predictive power of diagnostic markers for time-dependent disease outcomes. This analysis can compare multiple markers at the same time without the need to determine cutoff values. This is the first report in which the prognosis of PBC patients was assessed using continuous predictive markers and time-dependent ROC analysis.

This study clearly demonstrated that both the ALBI score and Mayo score, assessed at the start of follow-up, were strongly associated with outcomes (death or liver transplantation). Clinically, the C–P score/classification is commonly used to determine the prognosis of patients with chronic liver diseases or liver cancers. However, the C–P score is limited by its use of arbitrary parameter cutoff values and by the fact that all five parameters are weighted equally, including two subjective parameters: the presence of ascites and the degree of hepatic encephalopathy. By contrast, the ALBI score is calculated using only serum albumin and bilirubin levels, both of which are incorporated in the conventional C–P score, and it does not require evaluation of prothrombin time, another parameter in the C–P classification system. Recently, the ALBI score/grade has been widely used for predicting outcomes, especially in the field of HCC. Regarding the analysis of clinical outcomes in PBC, Fujita et al.
and Chan et al. reported the utility of the ALBI score as a predictive marker20,21. However, both studies used
the Kaplan–Meier method and/or Cox–proportional hazard model for outcome analysis, but neither statistical
technique can adequately analyze both short- and long-term outcomes. Since most prognostic markers for PBC
incorporate serum bilirubin and albumin levels, multivariate analysis is unsuitable for comparing these markers
because these levels can be confounding factors. A strength of our study is that it evaluates outcome data using
time-dependent ROC analysis comparing these studies at the same time points.

On the other hand, almost all prognostic markers used in the present study were selected as significant factors
for predicting not only liver-related but also non-liver related events by univariate analysis using Fine and Gray
proportional hazards models. In 15 cases died by non liver-related deaths, more than half patients (n = 8) died
due to non-malignant diseases: acute pneumonia; n = 4, gastrointestinal bleeding other than varices bleeding;
n = 3, acute pancreatitis; n = 1 (Supplementary Table S2. online). Especially, the incidence of pneumonia and GI
bleeding were related to the progression of PBC, and prognostic markers for PBC might be able to affect HRs
for these incidence.

Time-dependent ROC analysis showed that the AUROCs of ALBI score were higher than those of the Mayo
score within 2 years after the start of observation. Conversely, the Mayo score had a high AUROC throughout
the follow-up period. The Mayo score is one of the most widely accepted non-invasive markers for estimating the
prognosis of patients with PBC23. The high, stable AUROC of the Mayo score indicated its efficacy in predicting
both short- and long-term outcomes. However, one of its drawbacks is that like the C–P score, its calculation
requires the presence of edema, a subjective parameter. Therefore, the ALBI score/grade has advantages in terms
of simplicity and objectivity. Interestingly, the AUROC of the FIB-4 index, a non-invasive fibrosis marker used
widely in clinical settings, was the lowest of all the markers, although it gradually increased over time. To reveal
which factors in the component of FIB-4 index can be responsible for this change, we performed time-dependent
ROC analysis using each single marker (Supplementary Fig. S4). On the other hand, the FIB-4 index is calculated with
age; AST, ALT, and platelet counts, but this analysis revealed the age factor showed similar changes of the AUROCs
in FIB-4 index, indicating that this shift seems to be strongly influenced by age factor. These results indicate that
the FIB-4 index is not ideal for predicting prognosis in a short term. On the other hand, the AUROCs of total
bilirubin were the highest, and those of albumin were lowest among these single markers. Therefore, the high
AUROCs in ALBI score can be mainly affected by the bilirubin factor. However, the ALBI score that calculated
by bilirubin and albumin showed higher AUROCs than bilirubin alone, suggesting that it could be used as a
better prognostic marker by combining them. Previous reports showed that non-invasive fibrosis markers are
essential for estimating outcomes in PBC25,26. The ALBI and Mayo scores, which include bilirubin, are superior
to the FIB-4 index for predicting outcomes, including mortality and liver transplant-free survival, throughout
the clinical course of patients with PBC.

This study has several limitations. First, this was a retrospective, hospital-based cohort study, which increases
the risk of selection bias. We included patients with PBC from three clinical hospitals, but liver transplantation
was performed only in Nagoya University Hospital. Since accessibility to liver transplantation was limited in
the other two participated hospitals, there was a concern for analyzing outcomes. However, the ALBI score/grade
showed identical prognostic results between these hospitals (data not shown) implying the accessibility
did not affect the prognosis in PBC patients in our clinical settings. Second, although recent studies focusing
on transplant-free survival after 1 year of UDCA therapy demonstrated the utility of two additional prognostic
scores, the GLOBE score and UK-PBC score25,26, we were unable to assess prognosis using these scores. In this
study, almost all patients (92.3%) underwent treatment with UDCA during follow-up, but we did not assess the
response to UDCA. Our study included cases with rapid progression of liver dysfunction, and we investigated
not only long-term but also short-term prognosis, i.e., within 1 year. However, the GLOBE and UK-PBC scores
require data obtained at least 1 year after the start of UDCA therapy. In addition, 41 patients (10.0%) who have
already started the UDCA treatment at the start of follow-up. To assess several additional markers, including
the GLOBE and UK-PBC scores, further studies should recruit additional UDCA naïve patients with observation
periods over 1 year. Third, histological confirmations of liver fibrosis were obtained in only some patients. In
addition, in many cases there was a time lag from the start of follow-up to liver biopsy. Therefore, in this study
we did not assess the correlation between outcome and pathological liver fibrosis. However, the strengths of our
study included its long-term follow-up period and large number of patients.

In conclusion, our study showed that the ALBI score and Mayo score had high prognostic ability for predicting
outcomes in PBC patients. An advantage of the ALBI score/grade is that it can be simply calculated by objective
blood tests without subjective factors or invasive procedures. Therefore, the ALBI score/grade may be useful
to determine treatment strategies for PBC patients. Conversely, the FIB-4 index has limited ability to predict
prognosis in the short term. Clinicians following patients with PBC should be familiar with the characteristics
of each of these predictive markers.

Methods
Patients. Between March 1990 and October 2018, a total of 435 patients were diagnosed with PBC at Nagoya
University Hospital, Ogaki Municipal Hospital, and Toyota Kosei Hospital. The diagnosis of PBC was based on
criteria by the Japan Society of Hepatology16, as follows: Patients with one of the following criteria should be
diagnosed with PBC: (1) histologically confirmed chronic non-suppurative destructive cholangitis (CNSDC)
with laboratory findings compatible with PBC; (2) positivity for anti-mitochondrial antibodies (AMAs) with
histological findings compatible with PBC but in the absence of characteristic histological findings of CNSDC;
and (3) no histological findings available, but positivity for AMA as well as clinical findings and a course indica-
tive of typical cholestatic PBC. The pathological stage of PBC was evaluated by an experienced pathologist who
specialized in liver pathology based on the Scheuer’s classification as follows: stage 1, florid duct lesion; stage 2, ductular proliferation; stage 3, scarring; and stage 4, cirrhosis.

Twenty-six patients were excluded for the following reasons: (1) insufficient follow-up or incomplete data, n = 2; (2) PBC-AIH (autoimmune hepatitis) overlap syndrome, n = 15; (3) suspected liver injury induced by concomitant nonalcoholic steatohepatitis from hepatic pathology, n = 1; (4) concomitant hepatitis B and/or C virus infection, n = 3; (5) hepatocellular carcinoma (HCC) that did not meet the Milan Criteria, n = 5. Ultimately, 409 patients were enrolled in this retrospective study. We defined probable PBC-AIH overlap syndrome based on the combination of immunoglobulin G levels more than twice the upper limit of normal (ULN) and aminotransferase levels more than five times the ULN. If we suspected overlap syndrome, the final diagnosis was performed based on pathological findings on liver biopsy. Additionally, we excluded the patients who could improve their liver-related laboratory data by the administration of corticosteroids during the follow-up period as PBC-AIH overlap syndrome. In this study, for assessing liver transplantation-free survival in patients who met the Milan Criteria, we included patients with HCC who met the Milan Criteria at the time of diagnosis of PBC, as well as those with a history of complete cure of HCC. Regular surveillance was performed every 3–6 months using ultrasonography and/or blood tests, and included measurement of the tumor marker alpha-fetoprotein. Decisions regarding the treatment of each patient were based on Japanese treatment guidelines for PBC. The study protocol was approved by the institutional review board of Nagoya University Hospital (No. 2019-0055) and was in compliance with the Helsinki Declaration. Informed consent in the present study was obtained in the form of opt-out on the website of Nagoya University Hospital, Ogaki Municipal Hospital, and Toyota Kosei Hospital.

**Calculation of prognostic parameters.** The formulas of the MELD score, Mayo risk score for PBC, Fibrosis-4 (FIB-4) index, and ALBI score/grade were as follows:

- **MELD score** = \(0.957 \times \text{ln} (\text{serum creatinine}) + 0.378 \times \text{ln} (\text{serum bilirubin}) + 1.120 \times \text{ln} (\text{INR}) + 0.643 \times 10\); if on hemodialysis: calculate as creatinine 4.0 mg/dl.

- **Mayo risk score** = \(0.051 \times \text{age} + 1.209 \times \text{log} (\text{bilirubin}) - 3.304 \times \text{log} (\text{albumin}) + 2.754 \times \text{log} (\text{prothrombin time sec}) + 0.675 \times \text{edema}; \text{edema: 0 = no edema without diuretics, 0.5 = edema without diuretic therapy or edema resolved with diuretic therapy, 1 = edema despite diuretic therapy}\).

- **FIB-4 index** = \( \text{aspartate aminotransferase (AST)} \times \text{platelet count} \times (10^9/\text{L}) \times \text{alanine aminotransferase (ALT)} \times \text{serum bilirubin (µmol/L)} \times 0.66\).

- **ALBI score/grade** = \(\log_{10} \text{bilirubin (µmol/L)} \times 0.66 + \text{albumin (g/L)} \times -0.085\). Grade 1 liver function corresponds to an ALBI score \(\leq -2.60\), grade 2 corresponds to an ALBI score from \(-2.60\) to \(-1.39\), and grade 3 corresponds to an ALBI score \(> -1.39\).

In this study, we used the modified C–P score for PBC based on a previous report. These prognostic markers were calculated at the start of follow-up at each institution and used for analysis of the prognosis in the present study.

**Statistical analysis.** Continuous variables are expressed as mean ± SD (standard deviation). Categorical variables are expressed as number (percentage). Actuarial analysis of cumulative survival was performed using the Kaplan–Meier method based on C–P and ALBI grades, and differences were tested using the log-rank test. Discriminatory abilities of the scoring models were assessed using Aikake's information criterion (AIC). Fine-Grey proportional hazards models were used to calculate hazard ratios (HRs) for overall disease-related and liver transplantation-related mortality. Statistical significance was defined as \(P < 0.05\). Statistical analyses other than time-dependent ROC were performed with EZR (Saitama Medical Center, Jichi Medical University, Saitama, Japan), a graphical user interface for R (The R Foundation for Statistical Computing, Vienna, Austria). More precisely, it is a modified version of the R commander designed to add statistical functions frequently used in biostatistics. We used the “survivalROC” package, written for R, to assess marker performance using time-dependent ROC curves.

Received: 29 July 2020; Accepted: 6 October 2020
Published online: 20 October 2020

**References**

1. Lindor, K. D., Bowls, C. L., Boyer, I., Levy, C. & Mayo, M. Primary biliary cholangitis: 2018 practice guidance from the American Association for the Study of Liver Diseases. *Hepatology* 69, 394–419. https://doi.org/10.1002/hep.3145 (2019).

2. Selmi, C., Bowls, C. L., Gershwin, M. E. & Coppell, R. L. Primary biliary cirrhosis. *Lancet* 377, 1600–1609. https://doi.org/10.1016/S0140-6736(10)61965-4 (2011).

3. Zhang, L. N. et al. Early biochemical response to ursodeoxycholic acid and long-term prognosis of primary biliary cirrhosis: results of a 14-year cohort study. *Hepatology* 58, 264–272. https://doi.org/10.1002/hep.26322 (2013).

4. Pares, A., Caballería, L. & Rodes, J. Excellent long-term survival in patients with primary biliary cirrhosis and biochemical response to ursodeoxycholic acid. *Gastroenterology* 130, 715–720. https://doi.org/10.1053/j.gastro.2005.12.029 (2006).

5. Kamath, P. S. et al. A model to predict survival in patients with end-stage liver disease. *Hepatology* 33, 464–470. https://doi.org/10.1053/jhep.2001.22172 (2001).

6. van Dam, G. M. et al. Primary biliary cirrhosis: Dutch application of the Mayo Model before and after orthotopic liver transplantation. *HepatoGastroenterology* 44, 732–743 (1997).
9. Pugh, R. N., Murray-Lyon, I. M., Dawson, J. L., Pietroni, M. C. & Williams, R. Transection of the oesophagus for bleeding oesophageal varices. Br. J. Surg. 60, 646–649. https://doi.org/10.1002/bjs.1800600187 (1973).

10. Johnson, P. J. et al. Assessment of liver function in patients with hepatocellular carcinoma: a new evidence-based approach—the ALBI grade. J. Clin. Oncol. 33, 550–558. https://doi.org/10.1200/JCO.2014.57.9151 (2015).

11. Goet, J. C., Harms, M. H., Carbone, M. & Hansen, B. E. Risk stratification and prognostic modelling in primary biliary cholangitis. Best Pract. Res. Clin. Gastroenterol. 34–35, 95–106. https://doi.org/10.1016/j.bpg.2018.06.006 (2018).

12. Sterling, R. K. et al. Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. Hepatology 43, 1317–1325. https://doi.org/10.1002/hep.21178 (2006).

13. Heagerty, P. J. & Zheng, Y. Survival model predictive accuracy and ROC curves. Biometrics 61, 92–105. https://doi.org/10.1111/j.0006-341X.2005.03081.x (2005).

14. Heagerty, P. J., Lumley, T. & Pepe, M. S. Survival model predictive accuracy and ROC curves. Biometrics 61, 92–105. https://doi.org/10.1111/j.0006-341X.2005.03081.x (2005).

15. Marrero, J. A. et al. Diagnosis, staging, and management of hepatocellular carcinoma: 2018 practice guidance by the American Association for the Study of Liver Diseases. Hepatology 68, 723–750. https://doi.org/10.1002/hep.29913 (2018).

16. Working Subgroup for Clinical Practice Guidelines for Primary Biliary. C. Guidelines for the management of primary biliary cirrhosis: The Intractable Hepatobiliary Disease Study Group supported by the Ministry of Health, Labour and Welfare of Japan. Hepatol Res 44 Suppl 51, 71–90. https://doi.org/10.1111/hepr.12270 (2014).

17. Hiraoka, A. et al. Hepatic function during repeated TACE procedures and prognosis after introducing sorafenib in patients with unresectable hepatocellular carcinoma: multicenter analysis. Dig. Dis. 35, 602–610. https://doi.org/10.11159/000400226 (2017).

18. Hiraoka, A., Kumada, T., Michitaka, K. & Kudo, M. Newly proposed ALBI grade and ALBI-T score as tools for assessment of hepatic function and prognosis in hepatocellular carcinoma patients. Liver Cancer 8, 312–325. https://doi.org/10.1159/000494844 (2019).

19. Ho, S. Y. et al. Albumin-bilirubin grade-based nomogram of the BCLC system for personalized prognostic prediction in hepatocellular carcinoma. Liver Int. 40, 205–214. https://doi.org/10.1111/liv.14249 (2020).

20. Fujita, K. et al. Prediction of transplant-free survival through albumin–bilirubin score in primary biliary cholangitis. J. Clin. Med. https://doi.org/10.3390/jcm0801.258 (2019).

21. Chan, A. W. et al. New simple prognostic score for primary biliary cirrhosis: albumin–bilirubin score. J. Gastroenterol. Hepatol. 30, 1391–1396. https://doi.org/10.1111/jgh.12938 (2015).

22. Kim, W. R. et al. Adaptation of the Mayo primary biliary cirrhosis natural history model for application in liver transplant candidates. Liver Transpl 6, 489–494. https://doi.org/10.1053/jlts.2000.6503 (2000).

23. Murillo Perez, C. E. et al. Fibrosis stage is an independent predictor of outcome in primary biliary cholangitis despite biochemical treatment response. Aliment Pharmacol. Ther. 50, 1127–1136. https://doi.org/10.1111/apt.15533 (2019).

24. Joshita, S. et al. Clinical utility of FibroScan as a non-invasive diagnostic test for primary biliary cholangitis. J. Gastroenterol. Hepatol. https://doi.org/10.1111/jgh.14929 (2019).

25. Carbone, M. et al. The UK-PBC risk scores: derivation and validation of a scoring system for long-term prediction of end-stage liver disease in primary biliary cholangitis. Hepatology 63, 930–950. https://doi.org/10.1002/hep.28017 (2016).

26. Lammers, W. J. et al. Development and validation of a scoring system to predict outcomes of patients with primary biliary cirrhosis receiving ursodeoxycholic acid therapy. Gastroenterology 149, 1804–1812. https://doi.org/10.1053/j.gastro.2015.07.061 (2015).

27. Scheuer, P. Primary biliary cirrhosis. Proc. R. Soc. Med. 60, 1257–1260 (1967).

28. Mazaferro, V. et al. Liver transplantation for the treatment of small hepatocellular carcinomas in patients with cirrhosis. N. Engl. J. Med. 334, 693–699. https://doi.org/10.1056/NEJM199603143341104 (1996).

29. Han, Y. & Chen, Y. An interpretation of 2017 EASL clinical practice guidelines: the diagnosis and management of patients with primary biliary cholangitis. Zhonghua Gan Zang Bing Za Zhi 25, 814–818. https://doi.org/10.3760/cma.j.issn.1007-3418.2017.11.004 (2017).

30. Chazouilleres, O. et al. Primary biliary cirrhosis-autoimmune hepatitis overlap syndrome: clinical features and response to therapy. Hepatology 28, 296–301. https://doi.org/10.1002/hep.10280203 (1998).

31. Kanda, Y. Investigation of the freely available easy-to-use software “EZK” for medical statistics. Bone Marrow Transpl. 48, 452–458. https://doi.org/10.1038/bmt.2012.244 (2013).

Author contributions
Concept and study design: T.I., acquisition of data: T.I., D.N., S.Y., T.K., Drafting of the manuscript: T.I., M.F., Critical revision of the manuscript for important intellectual content: M.I., K.Y., N.I., Y.I., T.H., T.K., M.F., statistical analysis: T.I., H.M., T.T. Review and approval: all authors.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-74732-3.

Correspondence and requests for materials should be addressed to T.I.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
