npg

Development and Validation of a Novel Fibrosis Marker in Biliary Atresia during Infancy

Hirofumi Tomita, MD, PhD[1,2], Yasushi Fuchimoto, MD, PhD[3], Akihiro Fujino, MD, PhD[1], Ken Hoshino, MD, PhD[1], Yohei Yamada, MD[1], Yohei Masugi, MD, PhD[4], Michiie Sakamoto, MD, PhD[4], Mureo Kasahara, MD, PhD[5], Yutaka Kanamori, MD, PhD[3], Atsuko Nakazawa, MD, PhD[6], Fumiko Yoshida, MD[2], Seiya Akatsuka, MD, PhD[7], Miwako Nakano, MD, PhD[2] and Tatsuo Kuroda, MD, PhD[1]

OBJECTIVES: Most biliary atresia (BA) patients suffer from liver fibrosis and often require liver transplantation. The aim of this study was to develop and validate a novel fibrosis marker for BA patients aged < 1 year—the infant BA liver fibrosis (iBALF) score—subsequent to the previously reported fibrosis marker for BA patients aged ≥ 1 year.
METHODS: From three institutions for pediatric surgery, BA patients and their native liver histology examinations performed at the age of < 1 year were retrospectively identified and assigned to a development cohort (58 patients and 73 examinations) or validation cohort (92 patients and 117 examinations) according to their institutions. Histological fibrosis stages (F0–F4), blood test results, and clinical information at the time of liver histology examination were reviewed. The iBALF score was determined using multivariate ordered logistic regression analysis and was assessed for its associations with histological fibrosis stages.
RESULTS: The iBALF score equation was composed of natural logarithms, including serum total bilirubin level, blood platelet counts, and days of age. The score revealed a strong correlation with fibrosis stage ($r = 0.80$ and 0.73 in the development and validation cohorts, respectively; $P < 0.001$). The areas under the receiver-operating characteristic curves for diagnosing each fibrosis stage were 0.86–0.94 in the development cohort and 0.86–0.90 in the validation cohort ($P < 0.001$), indicating good diagnostic power. In addition, no patient with an iBALF score > 6 (equivalent to F4) at the initial surgery survived with their native liver at 1 year of age ($n = 9$).
CONCLUSIONS: The iBALF score that was developed was a good noninvasive marker of native liver fibrosis for BA patients aged < 1 year.

INTRODUCTION

Biliary atresia (BA) is a common cause of pediatric cholestasis due to obliterative cholangiopathy that develops in 1/5,000–1/19,000 newborns and is the most common indication for pediatric liver transplantation.[1] Because rapid progression of liver fibrosis is a prominent feature of BA patients, early diagnosis and timely surgical correction of cholestasis are needed.[1,2] In general, hepatoportoenterostomy is initially attempted to achieve initial bile drainage for most patients in whom the disease involves the bile duct at the porta hepatis (type 3 disease) and for whom a surgical anastomosis between the bile duct and the gastrointestinal tract cannot be created.[1] Although hepatoportoenterostomy can achieve initial bile drainage in 50–60% of cases, advanced liver fibrosis and possible progression of liver fibrosis after surgery lead to portal hypertension and cirrhosis.[1,2] Liver transplantation is performed secondarily when bile drainage is not achieved or when cirrhotic complications affect patients.[3] Thus, liver fibrosis is thought to be an important predictor of

outcome for BA patients, for whom long-term survival with the native liver is only achieved in ~20%.[2,3]

Although assessment of liver fibrosis is considered to be useful in BA patients, liver histology examinations are generally performed only at the same time as surgical procedures; liver tissue is obtained via surgical wedge biopsy during laparotomy or total hepatectomy during liver transplant surgery; postsurgical liver biopsy examinations for monitoring fibrosis progression are not generally performed.[2] However, we have performed postsurgical liver biopsy examinations to more precisely evaluate native liver status and to determine the optimal timing for liver transplantation, mostly from living donors in Japan, in clinical practice. Because reliable, surrogate, noninvasive liver fibrosis markers in BA patients have been limited,[2] we previously developed a BA liver fibrosis (BALF) score using a retrospective analysis of postsurgical native liver histology examinations.[4] The BALF score was calculated using standard liver test results and age and is a potential liver fibrosis marker in BA patients aged ≥ 1 year; however, the score was unable to predict liver fibrosis in

[1]Department of Pediatric Surgery, Keio University School of Medicine, Tokyo, Japan; [2]Department of Pediatric Surgery, Saitama City Hospital, Midori-ku, Saitama-shi, Japan; [3]Department of Surgery, National Center for Child Health and Development, Okura Setagaya-ku Tokyo, Japan; [4]Department of Pathology, Keio University School of Medicine, Shinjuku-ku, Japan; [5]Department of Transplant Surgery, National Center for Child Health and Development, Okura Setagaya-ku Tokyo, Japan; [6]Department of Pathology, National Center for Child Health and Development, Okura Setagaya-ku Tokyo, Japan and [7]Department of Pathology, Saitama City Hospital, Midori-ku, Saitama-shi, Japan
Correspondence: Tatsuo Kuroda, MD, PhD, Department of Pediatric Surgery, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan.
E-mail: kuroda-t@z8.keio.jp
Received 25 June 2015; accepted 12 October 2015

patients aged <1 year.[4] Because some patients require primary or early liver transplantation owing to rapid progression of liver fibrosis, we considered that an available, reliable and noninvasive liver fibrosis marker during infancy would be of great worth. In the current study, we developed a novel noninvasive fibrosis marker for BA patients aged <1 year, subsequent to the previously reported BALF score. This novel fibrosis marker was delineated as the infant BALF (iBALF) score and was validated in an independent population of BA patients.

METHODS

Study population and ethical considerations. The medical records of BA patients at three institutions for pediatric surgery were retrospectively reviewed, and 155 patients from whom native liver specimens had been obtained at <1 year of age between March 1993 and April 2014 were identified. The patients were assigned to either the development cohort ($n=60$) or the validation cohort ($n=95$), according to the participating institutions: the development cohort derived from Keio University Hospital and Saitama City Hospital, and the validation cohort derived from the National Center for Child Health and Development. We confirmed that the development and validation cohorts did not share the same patient. This study conformed to the ethical guidelines of the 1975 Declaration of Helsinki and was approved by the ethical committees of all three participating institutions. All of the biopsies and surgeries were performed after obtaining written informed consent.

Liver tissue sampling and histology examinations. During the initial bile drainage surgery, wedge biopsy examinations were performed using surgical resection from the edge of the liver. Postsurgical liver histology examinations were performed in several patients from wedge biopsy specimens during re-laparotomy and from percutaneous liver biopsy specimens of ≥ 1.0 cm in length using an 18-gauge suction needle under ultrasonographic guidance. Explanted livers were obtained during liver transplant surgery and were histologically examined. Histological liver fibrosis stages were based on the documented findings by experienced pathologists at the time liver tissue samples were obtained; if needed, re-evaluation by an experienced pathologist participating in the current study was performed at each institution. For liver fibrosis grading, the Metavir scoring system[5] or the new Inuyama classification[6] was used with the following classifications: F0, no portal fibrosis; F1, portal fibrosis without septa; F2, portal fibrosis with rare septa; F3, numerous septa or lobular distortion without cirrhosis; and F4, cirrhosis.

Data collection and data exclusion. The patients' clinical information and blood test results were collected from the medical records in association with liver histology examinations. The collected clinical information included sex, disease type, history of surgical procedure, age at the time of surgery, age at tissue sampling, and method of tissue sampling. Patients who had a history of splenectomy or partial splenic embolization and those with BA splenic malformation

syndrome were excluded. The disease type was determined according to the classification of the Japanese Biliary Atresia Society:[7] atresia at the level of the most proximal part of the common bile duct (type 1), hepatic duct (type 2), and porta hepatis (type 3). The collected blood test results included serum total bilirubin (TB), direct bilirubin, aspartate aminotransferase, alanine aminotransferase, γ-glutamyltransferase (GGT), albumin, and cholinesterase levels; prothrombin time-international normalized ratios; and platelet counts, which had been examined within a few days before liver tissue sampling. The impact of transfusion, cholangitis, and vitamin K deficiency on the blood test results was excluded to the greatest extent possible; if transfusion had been performed or cholangitis had occurred before liver tissue sampling, data preceding transfusion or cholangitis up to 1 month were used, whereas in cases of vitamin K deficiency at the time of initial surgery, data after correction of vitamin K deficiency were used. Cholangitis was defined as fever and serum TB elevation without any other apparent cause, and vitamin K deficiency was defined as coagulopathy that improved soon after vitamin K administration.

Development of the iBALF score. Development of the iBALF score was accomplished using a similar method to BALF score development.[4] To predict the histological fibrosis stage, ordered logistic regression analyses were performed, using the semiquantitative histological fibrosis grading as ordinal data (from F0 to F4) for the dependent variable; the logarithmic values of the collected blood test results and days of age at the time of corresponding histological examination served as the independent variables. To determine the iBALF score equation, significant independent variables and the regression coefficients from the multivariate analysis were used. The constant of the score equation was determined by bringing the cutoff values of the iBALF score for fibrosis prediction close to the previously reported BALF score cutoff values in patients aged ≥ 1 year (2.42 for $\geq$F2, 4.12 for $\geq$F3, and 5.64 for F4).[4]

Assessment of the iBALF score. After determination of the iBALF score equation from the development cohort, the scores were calculated from the development and validation cohort data; the values of the iBALF score were obtained along with the corresponding histological examination results. The diagnostic power of the iBALF score for predicting each fibrosis stage was assessed using a receiver-operating characteristic curve comparing the blood platelet counts and the aspartate aminotransferase-to-platelet ratio index (APRI), which has been the most widely investigated fibrosis marker in BA patients. The APRI was calculated using the following equation:[8]

APRI = (aspartate aminotransferase/upper normal limit/platelet counts (10^9/l)) × 100.

The upper normal limit of aspartate aminotransferase was determined according to the age-specific reference intervals for Japan.[9]

Assessment of the prognosis at 1 year of age. The prognosis of the patients who participated in the study from the initial surgery (initial bile drainage surgery or primary

liver transplantation) was assessed using serial data collection. The prognosis at 1 year of age was investigated as either death before liver transplantation, receiving liver transplantation before 1 year of age, or surviving with their native liver. Among the patients surviving with their native liver at 1 year of age, the earliest blood test results after reaching 1 year of age were collected from the medical records; if transfusion had been performed or cholangitis had occurred before the blood test was performed, the data at >1 month after transfusion or cholangitis were selected. The BALF score that had been developed to predict liver fibrosis stage in BA patients aged ≥ 1 year was then used to evaluate the status of the native liver. The BALF score was calculated using the following equation:[4]

BALF score = $7.196 + 1.438 \times Log_e$ [TB (mg/dl)] + $0.434 \times Log_e$ [GGT (IU/l)] − $3.491 \times Log_e$ [albumin (g/dl)] − $0.670 \times Log_e$ [age (years)].

Statistical analysis. The categorical and ordinal data are presented as frequencies and were statistically compared using the Fisher exact test. The continuous data are presented as medians (ranges) and were statistically compared using the Mann–Whitney U-test. Correlations between the ordinal and/or continuous data were assessed by the Spearman correlation coefficient (r). For logistic regression analyses, the P value of each independent variable was determined using the Wald χ^2-value (Wald), which was calculated by squaring the ratio of the regression coefficient divided by its standard error. For receiver-operating characteristic curve analyses, areas under the curve (AUCs) were calculated; an AUC of 1.0 indicates a test of perfect diagnostic power, whereas an AUC of 0.5 indicates no diagnostic power. Differences between AUCs were examined using the DeLong test. The cutoff values were determined as the points that showed high sensitivity and specificity in a balanced manner. P values <0.05 were considered statistically significant. Statistical analyses were performed using SPSS 22.0 software (IBM SPSS, Chicago, IL, USA) and R 3.1.0 software (The R Foundation for Statistical Computing Vienna, Austria; http://www.R-project.org/).

RESULTS

Patient characteristics. We excluded two and three patients with BA splenic malformation syndrome from the development and validation cohorts, respectively. No patient had a history of splenectomy or partial splenic embolization before data collection. One histology examination using percutaneous needle biopsy obtained after the initial surgery from a development cohort patient was inappropriate for evaluation and was excluded from the study. After exclusions, the development cohort included 58 patients and 73 liver histology examinations, and the validation cohort included 92 patients and 117 liver histology examinations. The timing of the patients' participation and tissue sampling in the development and validation cohorts is summarized in Figure 1. Patient characteristics according to the development and validation cohorts are shown in Table 1. Significant differences between the development and validation cohorts were found in the frequencies of disease type ($P=0.02$) and
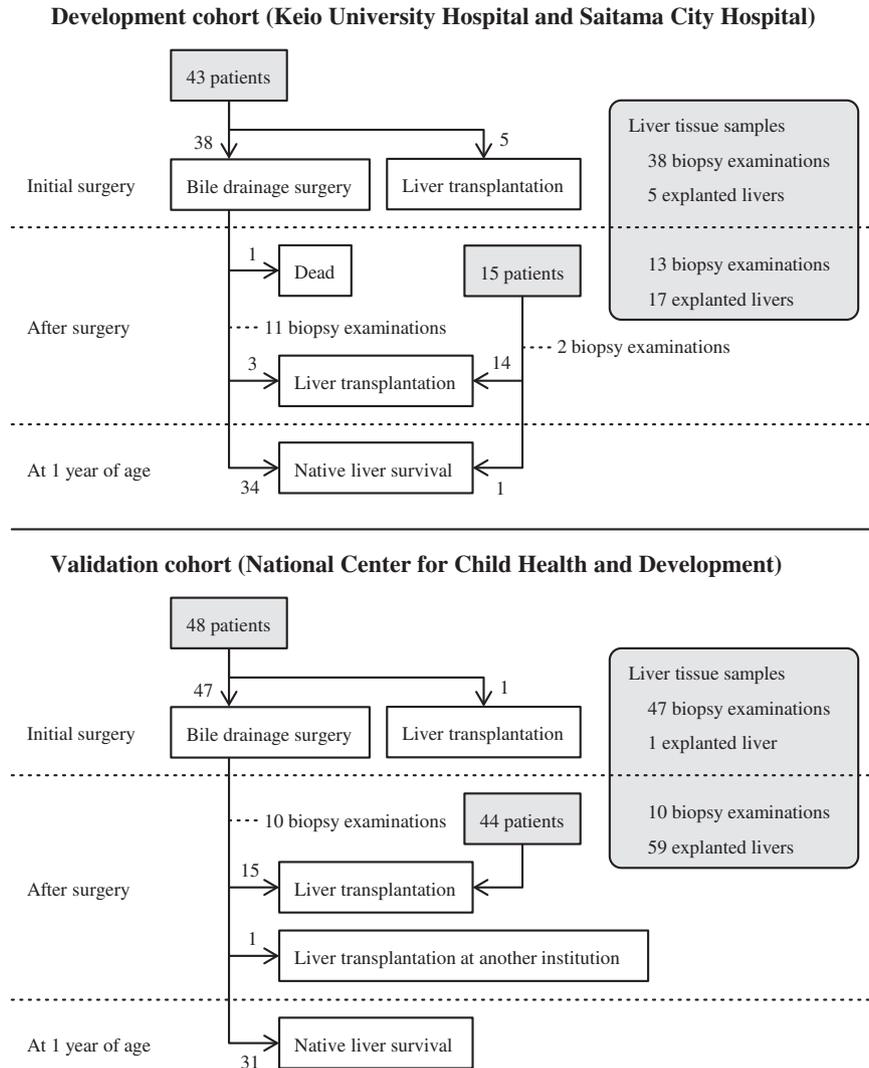
initial bile drainage surgical procedure ($P=0.03$): the validation cohort included more patients with type 3 disease requiring hepatoportoenterostomy. Significant differences regarding liver transplantation before 1 year of age were also found: the validation cohort included fewer patients received primary liver transplantation, and more patients received liver transplantation after bile drainage surgery than in the development cohort ($P<0.001$). Days of age at the time of liver transplantation were significantly lower in the validation cohort than in the development cohort ($P=0.009$).

Liver histology and blood test results. In the development cohort, 10 (13.7%) histology examinations showed a liver fibrosis stage of F1, whereas 19 (26.0%) showed a stage of F2, 20 (27.4%) showed a stage of F3, and 24 (32.9%) showed a stage of F4. In the validation cohort, eight (6.8%) histology examinations showed a stage of F1, 23 (19.7%) showed a stage of F2, 27 (23.1%) showed a stage of F3, and 59 (50.4%) showed a stage of F4. Liver histology examinations and the corresponding blood test results from the development and validation cohorts according to the biopsy examination or liver transplantation are presented in Table 2. At the time of biopsy examinations, serum direct bilirubin levels were significantly lower and serum albumin levels were significantly higher in the development cohort than in the validation cohort ($P=0.03$ and $P<0.001$, respectively), because the development cohort involved a greater number of needle biopsy examinations, which were performed for patients with a better surgical response than the validation cohort ($P=0.002$). At the time of liver transplantation, blood test results were significantly worse in the development cohort than in the validation cohort, indicating different timing of liver transplant surgery between the cohorts.

Determination of the iBALF score equation. The results of the ordered logistic regression analyses in the development cohort are shown in Table 3. In the univariate analyses, natural logarithms of the blood platelet counts provided the highest significance (Wald = 31.461, $P<0.001$). In the multivariate analysis, the second significant independent variable was identified as natural logarithms of the serum TB levels using a forward selection method. As the third independent variable, natural logarithms of the prothrombin time-international normalized ratios and days of age were significant; we selected the days of age, because the distribution of the iBALF score approached the distribution of the previously reported BALF score. Finally, natural logarithms of the serum TB levels, blood platelet counts, and days of age at examination were selected as significant independent variables. The iBALF score equation was determined as:

iBALF score = $8 + 1.185 \times Log_e$ [TB (mg/dl)] − $1.882 \times Log_e$ [platelet count (10^9/l)] + $1.093 \times Log_e$ [age (days)].

iBALF scores according to the liver fibrosis stages. Figure 2 shows the boxplots for the iBALF score and APRI vs. the histological fibrosis stages in the development and validation cohorts. The iBALF score was more strongly correlated with the histological fibrosis stage than the APRI in both cohorts ($r=0.80$ and 0.73 in the development and validation cohorts, respectively; $P<0.001$). Between the

Development cohort (Keio University Hospital and Saitama City Hospital)

Validation cohort (National Center for Child Health and Development)

Figure 1 Timing of the patients' participation and tissue sampling in the development and validation cohorts. The number was counted after excluding two and three patients with biliary atresia splenic malformation syndrome from the development and validation cohorts, respectively.

cohorts, the iBALF score in the histology examinations displaying F4 showed a significant difference ($P = 0.006$); the median iBALF score values were 8.08 (range, 4.75–10.71) in the development cohort and 6.84 (range, 2.88–9.69) in the validation cohort. No significant difference was found in the other histological fibrosis stage groups.

Diagnostic power of the iBALF score. Figure 3 shows the receiver-operating characteristic curves of the iBALF score for diagnosing each fibrosis stage, compared with the APRI. In the development cohort, the AUCs of the iBALF score were 0.84 for a fibrosis stage $\geq$ F2, 0.91 for $\geq$ F3, and 0.96 for F4 ($P < 0.001$). In the validation cohort, the AUCs of the iBALF score were 0.86 for $\geq$ F2, 0.90 for $\geq$ F3, and 0.89 for F4 ($P < 0.001$); the diagnostic power for F4 fibrosis appeared to be worse than in the development cohort. The AUCs of the iBALF score were significantly greater than those of the APRI in diagnosing $\geq$ F2 ($P = 0.03$) and F4 ($P = 0.01$) in the development cohort, indicating more favorable diagnostic

power than the APRI; no significant difference was found in diagnosing $\geq$ F3 in the development cohort and in diagnosing $\geq$ F2, $\geq$ F3, and F4 in the validation cohort.

Cutoff value and diagnostic accuracy of the iBALF score. The cutoff values and diagnostic accuracies of the iBALF score for predicting histological fibrosis stages are shown in Table 4. The cutoff values of the development cohort were 3.00 for a fibrosis stage $\geq$ F2, 3.99 for $\geq$ F3, and 5.75 for F4, which were brought close to the previously reported cutoff values of the BALF score by adjusting the constant of the iBALF score equation. The diagnostic accuracies of the iBALF score for each fibrosis stage were acceptable: 78.1–93.2% in the development cohort and 80.3–82.9% in the validation cohort. The validation cohort appeared to have lower diagnostic accuracy for F4 diagnosis than the development cohort (82.0% vs. 93.2%, respectively).

Table 1 Patient characteristics of the development and validation cohorts

	Development cohort	Validation cohort	*P*-value
Number of patients	58	92	
Sex (male/female)	25/33	28/64	0.12
Disease type (type 1/type 2/type 3/unknown)	9/2/45/2	6/0/85/1	0.02
Initial bile drainage surgery (hepaticoenterostomy/hepatoportoenterostomy/none)	3/50/5	2/89/1	0.03
Days of age at the initial bile drainage surgery	74 (17–151) (*n* = 53)	73 (27–195) (*n* = 91)	0.28
Liver transplantation before 1 year of age (primary/after bile drainage surgery/none)	5/17/36	1/60/31	<0.001
Days of age at liver transplantation before 1 year of age	290 (179–356) (*n* = 22)	233 (126–346) (*n* = 61)	0.009
Number of histology examinations per each patient (1/2/3/4)	46/10/1/1	69/21/2/0	0.59

The categorical and ordinal data are presented as the number of patients and were statistically compared using the Fisher exact test. The continuous data are presented as medians (ranges) and were statistically compared using the Mann–Whitney *U*-test.

Table 2 Comparisons of the liver histology examinations and corresponding blood test results between the development and validation cohorts according to the biopsy examination or liver transplantation

	Biopsy examination			Liver transplantation		
	Development cohort	Validation cohort	*P*-value	Development cohort	Validation cohort	*P*-value
Number of examinations	51	57		22	60	
Wedge/needle	41/10	56/1	0.002			
Fibrosis stage (F1/F2/F3/F4)	10/19/18/4	8/23/19/7	0.78	0/0/2/20	0/0/8/52	0.72
Days of age	79 (17–328)	77 (27–345)	0.96	290 (179–356)	232 (126–346)	0.01
Blood test results						
TB (mg/dl)	8.0 (0.4–14.5)	8.3 (0.6–25.8)	0.06	20.6 (5.5–47.7)	12.1 (1.2–33.9)	<0.001
DB (mg/dl)	4.9 (0.1–9.5)	5.6 (0.3–17.6)	0.03	14.5 (3.2–34.4)	8.7 (0.6–22.1)	0.001
AST (IU/l)	161 (35–917)	150 (44–473)	0.77	269 (55–560)	162 (61–659)	0.007
ALT (IU/l)	109 (15–922)	110 (24–447)	0.98	127 (30–240)	110 (29–426)	0.44
GGT (IU/l)	582 (62–3434)	741 (36–2610)	0.15	124 (50–1010)	253 (20–1452)	0.28
Albumin (g/dl)	3.9 (2.3–4.8)	3.6 (2.6–4.3)	<0.001	3.2 (2.2–4.1)	3.0 (1.9–4.2)	0.72
ChE (IU/l)	279 (116–461)	270 (128–395)	0.86	140 (53–334)	143 (57–367)	0.73
PT-INR	1.03 (0.84–1.48)	1.00 (0.81–1.91)	0.19	1.41 (0.95–2.54)	1.28 (0.95–2.18)	0.047
Platelet count (×10⁹/l)	448 (172–1092)	444 (111–982)	0.93	118 (48–276)	196 (34–760)	0.02

ALT, alanine aminotransferase; AST, aspartate aminotransferase; ChE, cholinesterase; DB, direct bilirubin; GGT, γ-glutamyltransferase; PT-INR, prothrombin time-international normalized ratio; TB, total bilirubin. The categorical and ordinal data are presented as the number of examinations and were statistically compared using the Fisher exact test. The continuous data are presented as medians (ranges) and were statistically compared using the Mann–Whitney *U*-test.

Prognosis at 1 year of age according to the iBALF score at the initial surgery. Figure 4 shows the relationships between the iBALF score at the initial surgery and outcomes. The outcomes are presented as the need for liver transplantation before 1 year of age or as the BALF score at 1 year of age as a noninvasive liver fibrosis marker. None of the nine patients with an iBALF score >6 survived with their native liver: five patients in the development cohort underwent liver transplantation as the initial surgery, and four patients in the validation cohort required liver transplantation before 1 year of age. Among the patients who survived with their native liver at 1 year of age, the correlations between the iBALF score at the initial surgery and the BALF score at 1 year of age were not significant in the development (*n* = 34, *r* = 0.19, *P* = 0.29) or validation (*n* = 31, *r* = 0.04, *P* = 0.81) cohorts.

DISCUSSION

The BALF score was the first noninvasive fibrosis marker developed specifically for postsurgical BA patients aged ≥ 1 year; herein, the iBALF score was additionally developed for BA patients aged < 1 year. Although the BALF score calculated for patients aged < 1 year was previously reported

to show apparently high values regardless of the liver fibrosis stages,[4] the iBALF score showed strong correlations with the histological liver fibrosis stages and good diagnostic powers for each fibrosis stage in the development and validation cohorts. The differences between the BALF and iBALF scores in patients aged < 1 year were mainly derived from serum GGT level (included in the BALF score) and age (included in both scores), both of which had reverse coefficients in the logistic regression analyses for predicting liver fibrosis stages. Serum GGT elevation was reported to be associated with advanced fibrosis in patients aged ≥ 1 year,[4] but the current study indicated that serum GGT elevation was associated with less-advanced fibrosis in patients aged < 1 year. The effects of age on liver fibrosis progression were positive in patients aged < 1 year and negative in patients aged ≥ 1 year.[4] Although different equations were needed, we adjusted the iBALF score to have similar values for each fibrosis stage as the previously reported BALF score values in patients aged ≥ 1 year, this will aid in more easily understanding the iBALF scores in comparison with BALF scores, regardless of the age of the child. We suggest that the iBALF and BALF scores can monitor liver fibrosis in a similar manner before and after 1 year of age, respectively.

Table 3 Ordered logistic regression analyses for predicting liver fibrosis stages in the development cohort

| Variable	Coefficient (95% confidence interval)	Standard error	Wald	*P*-value
Univariate analysis				
$\mathrm{Log_e}$ (platelet count ($\times 10^9$/l))	− 2.859 (−3.858 to − 1.860)	0.510	31.461	<0.001
$\mathrm{Log_e}$ (age (days))	1.812 (1.119–2.506)	0.354	26.213	<0.001
$\mathrm{Log_e}$ (TB (mg/dl))	1.517 (0.891–2.142)	0.319	22.565	<0.001
$\mathrm{Log_e}$ (albumin (g/dl))	− 7.950 (−11.270 to − 4.631)	1.694	22.038	<0.001
$\mathrm{Log_e}$ (PT-INR)	7.126 (4.125–10.127)	1.531	21.662	<0.001
$\mathrm{Log_e}$ (ChE (IU/l))	− 2.841 (−4.078 to − 1.604)	0.631	20.272	<0.001
$\mathrm{Log_e}$ (DB (mg/dl))	1.269 (0.706–1.832)	0.287	19.534	<0.001
$\mathrm{Log_e}$ (GGT (IU/l))	− 0.926 (−1.398 to − 0.454)	0.241	14.772	<0.001
$\mathrm{Log_e}$ (AST (IU/l))	0.924 (0.235–1.612)	0.351	6.920	0.009
$\mathrm{Log_e}$ (ALT (IU/l))	0.278 (−0.312–0.868)	0.301	0.852	0.36
Multivariate analysis				
$\mathrm{Log_e}$ (TB (mg/dl))	1.185 (0.574–1.796)	0.312	14.452	<0.001
$\mathrm{Log_e}$ (platelet count ($\times 10^9$/l))	− 1.882 (−3.052 to − 0.712)	0.597	9.935	0.002
$\mathrm{Log_e}$ (age (days))	1.093 (0.232–1.955)	0.439	6.190	0.01

ALT, alanine aminotransferase; AST, aspartate aminotransferase; ChE, cholinesterase; DB, direct bilirubin; GGT, γ-glutamyltransferase; PT-INR, prothrombin time-international normalized ratio; TB, total bilirubin.
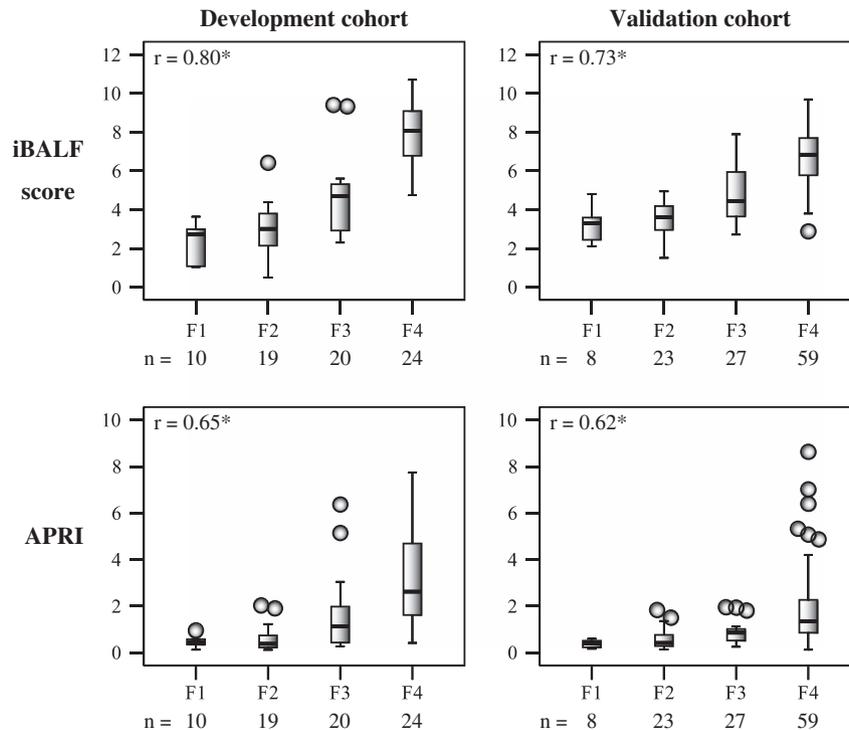
Figure 2 Values of the infant biliary atresia liver fibrosis (iBALF) score and aspartate aminotransferase-to-platelet ratio index (APRI) according to the histological fibrosis stages. Boxplots show the median values with the interquartile ranges, and error bars indicate the smallest and the largest values within 1.5 box-lengths of the upper and the lower quartiles. Circles represent the individual points for outliers. Correlations between the markers and the fibrosis stages were evaluated using the Spearman correlation coefficient (*r*); *P* < 0.001.

For infants with BA at presentation, two types of surgical procedure could be chosen—bile drainage surgery or liver transplantation. There were two reports regarding effects on outcomes after liver transplantation comparing early failure of hepatoportoenterostomy, which was defined as the need for liver transplantation within the first year of life, and primary liver transplantation. Alexopoulos *et al.*[10] described that early failure of hepatoportoenterostomy adversely affected patient and graft survival rates. Neto *et al.*[11] reported that early failure

of hepatoportoenterostomy had no effect on patient and graft survival, that late failure of hepatoportoenterostomy had a protective effect compared with primary liver transplantation, and that previous hepatoportoenterostomy increased biliary complications and bowel perforations after liver transplantation. Thus, it is important to know which patients can benefit from bile drainage surgery at presentation. In this study, we attempted to reveal the association between the iBALF score at the initial surgery and prognosis using the BALF score at
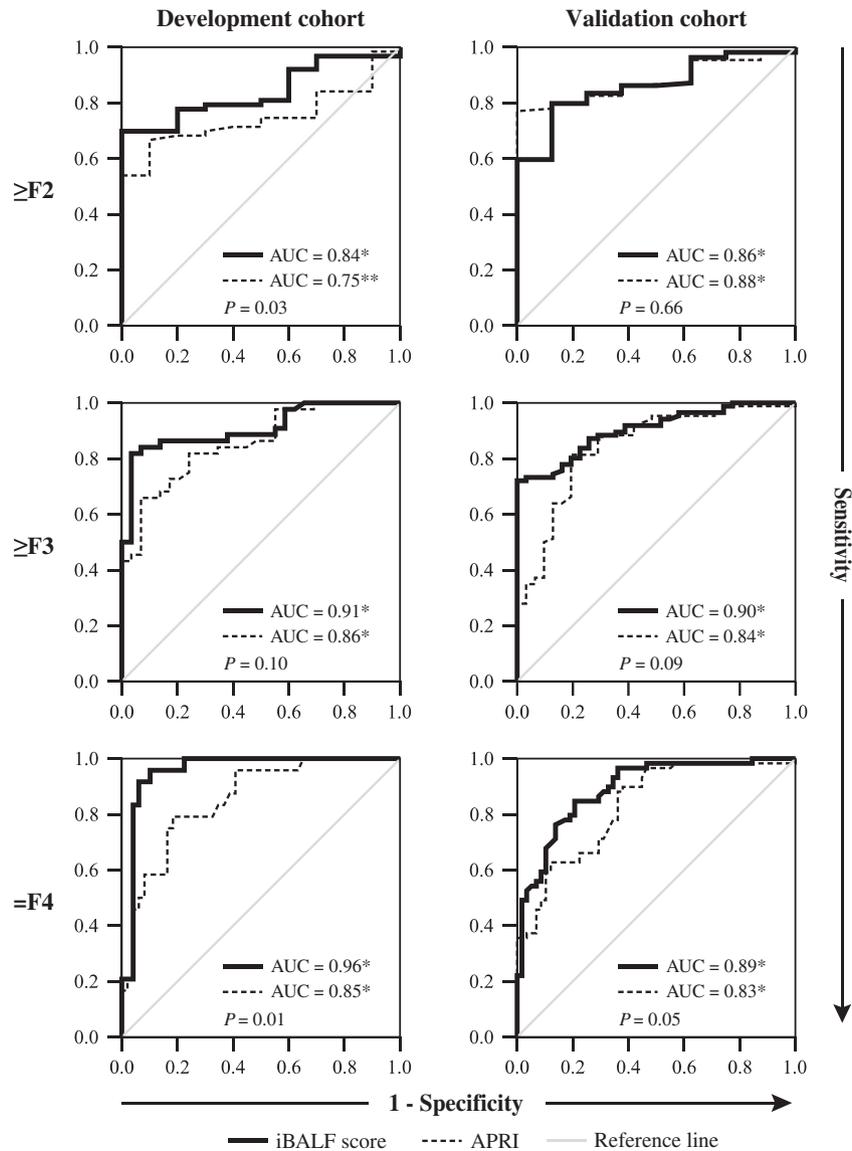
Figure 3 Receiver-operating characteristic curves of two fibrosis markers for diagnosing each fibrosis stage. Evaluated noninvasive markers included the infant biliary atresia liver fibrosis (iBALF) score (thick lines) and the aspartate aminotransferase-to-platelet ratio index (APRI, dashed lines). Gray lines indicate the reference lines. The diagnostic power of each marker was assessed by calculating the area under the curve (AUC); *$P < 0.001$, **$P = 0.01$. The P values in the panels represent the differences between AUCs of the iBALF score and the APRI using the DeLong test.

1 year of age. The results (Figure 4) suggest that BA patients with an iBALF score >6 at presentation might require liver transplantation rather than bile drainage surgery. However, the number of these severely affected patients was small in both cohorts. Except for these severely affected patients, the iBALF score at the initial surgery did not seem to be associated with native liver survival at 1 year of age. There was no correlation between the iBALF score at the initial surgery and the BALF score at 1 year of age among the patients with native liver survival, suggesting that liver fibrosis at the initial surgery had a limited effect on liver fibrosis progression or remission. We previously reported similar data on the actual fibrosis stages in 15 patients aged ≥ 2 years who underwent serial histological examinations at the time of initial surgery and after surgery and

who were included in the development cohort of the current study: seven of these 15 patients showed remission of fibrosis, five showed the same fibrosis stage, and three showed progression of fibrosis.[4] We believe that effective postsurgical antifibrotic therapy for BA patients is needed and that noninvasive fibrosis monitoring would be highly valuable in clinical practice and study.

In addition to our previous report, several other studies have proposed noninvasive markers to assess liver fibrosis in BA patients. The APRI, which was originally developed to predict cirrhosis in hepatitis C patients,[8] has been widely investigated in BA patients. Kim *et al.*[12] described that the correlation coefficient between the APRI and Metavir fibrosis score from 35 patients at the time of hepatoportoenterostomy was

0.77 ($P < 0.001$) and that the AUCs of the APRI for $\geq$ F3 and F4 fibrosis were 0.92 and 0.91, respectively. By contrast, Lind *et al.*[13] reported that the APRI was not significantly different according to the fibrosis stage in 31 patients at the time of hepatoportoenterostomy. In 23 patients after successful hepatoportoenterostomy (median, 4.2 years; range,

1.6–18.9 years after surgery), Lampela *et al.*[14] described a significant correlation between the APRI and Metavir fibrosis score ($r = 0.63$, $P < 0.001$) and a good diagnostic accuracy of the APRI for $\geq$ F3 with 93% sensitivity and 67% specificity. Another noninvasive fibrosis marker, transient elastography (Fibroscan), was more recently investigated to assess liver

Table 4 Cutoff values and diagnostic accuracies of the infant biliary atresia liver fibrosis (iBALF) score for predicting histological fibrosis stages

	n (%)	Cutoff	Sensitivity	Specificity	Accuracy
Development cohort (n = 73)					
$\geq$ F2	63 (86.3%)	3.00	77.8%	80.0%	78.1%
$\geq$ F3	44 (60.3%)	3.99	86.4%	86.2%	86.3%
= F4	24 (32.9%)	5.75	91.7%	93.9%	93.2%
Validation cohort (n = 117)					
$\geq$ F2	109 (93.2%)	3.56	83.5%	75.0%	82.9%
$\geq$ F3	86 (73.5%)	4.34	80.2%	80.6%	80.3%
= F4	59 (50.4%)	5.12	84.7%	79.3%	82.0%
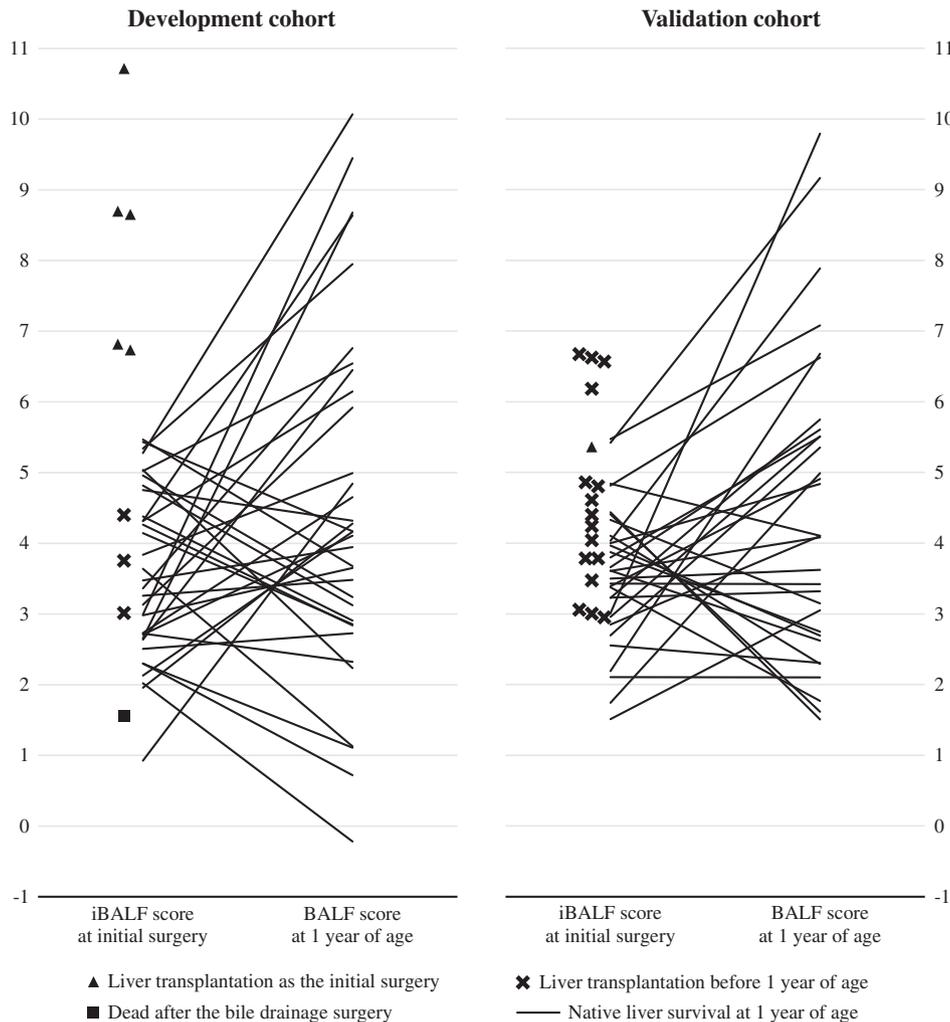
Figure 4 Relationships between the infant biliary atresia liver fibrosis (iBALF) score at the initial surgery and prognosis. Triangles indicate the patients receiving liver transplantation as the initial surgery. Crosses represent the patients requiring liver transplantation after bile drainage surgery before 1 year of age. The square indicates the patient who died after bile drainage surgery. The patients who survived with their native liver at 1 year of age are expressed by lines between the iBALF score at the bile drainage surgery and the biliary atresia liver fibrosis (BALF) score at 1 year of age.

stiffness using the ultrasound technique; Shin *et al.*[15] described that liver stiffness measurements obtained via transient elastography significantly correlated with Metavir fibrosis stages ($r = 0.63$, $P < 0.001$) and had good diagnostic powers for predicting severe fibrosis ($\geq$ F3; AUC = 0.86) and cirrhosis (F4; AUC = 0.96) in 47 BA patients aged < 1 year at the time of hepatoportoenterostomy with liver biopsy or liver transplantation. Moreover, the APRI and transient elastography had already been investigated for associations with esophageal varices, an important consequence of liver fibrosis and portal hypertension, in postsurgical BA patients.[14,16–18] The current study suggests the advantages of the iBALF score over the APRI: stronger correlation with the fibrosis stages and more favorable diagnostic power than the APRI. Unlike the elastography methods, the iBALF score has good accessibilities, such as no need for a special device and simple equation components that allow retrospective calculation.

Although the current study indicated that the iBALF was a good noninvasive fibrosis marker even in the validation cohort, it has several limitations. First, patients were selected from three institutions, two of which were assigned to the development cohort and one to the validation cohort, resulting in significant differences in patient characteristics and blood test results between the cohorts. BA patients aged < 1 year can be divided into three situations: patients before surgery, patients with a good postsurgical course, and patients requiring liver transplantation after bile drainage surgery. Although we intended that the iBALF-scoring system could apply in all situations, needle biopsy examinations for postsurgical patients with good bile drainage were performed at only one of the three participating institutions, thus the sample size was too small. To reflect the data from patients with a good postsurgical course in the iBALF score composition, we assigned the small number of these patients to the development cohort rather than randomly assigning them to the development cohort or the validation cohort. Thus, the relationships between liver fibrosis stage and the iBALF score of patients with a good postsurgical course could not be validated. In addition, there was a probable difference in the timing of liver transplantation between the institutions. Because of serious deceased donor organ shortages in Japan,[19] the timing of liver transplantation using liver allografts from living donors probably reflected the transplantation policy of each institution, resulting in significantly different ranges of the iBALF score in F4 patients between the cohorts and wide overlap in the ranges of the F3 and F4 groups in the validation cohort. The second limitation was general problems in prior studies of noninvasive fibrosis markers using the biopsy examinations as a reference standard: namely, biopsy sampling errors,[20] and observer variability.[21] Subcapsular wedge biopsy examination, which was used in most subjects in the current study, would tend to overestimate liver fibrosis. Thus, the fibrosis stages evaluated based on liver biopsy examinations might have false-positive and false-negative results.

In this study, we developed the iBALF score as a noninvasive surrogate fibrosis marker for BA patients aged < 1 year, in addition to the previously developed BALF-scoring system for BA patients aged $\geq$ 1 year. Although some concerns remain, the iBALF score was validated to strongly correlate with liver fibrosis stage and to have good diagnostic powers for predicting liver fibrosis. The iBALF and BALF scores may be useful in future clinical studies as surrogate fibrosis markers.

CONFLICT OF INTEREST

Study Highlights

WHAT IS CURRENT KNOWLEDGE

✓ Although liver fibrosis is a prominent feature of biliary atresia (BA) patients, noninvasive liver fibrosis markers in BA patients have been limited.

✓ We previously developed a BA liver fibrosis (BALF) score as the first specific liver fibrosis marker for BA patients aged $\geq$ 1 year.

WHAT IS NEW HERE

✓ We developed a novel noninvasive fibrosis marker for BA patients aged < 1 year—the infant BALF (iBALF) score.

✓ The iBALF score was validated to be a good noninvasive marker of native liver fibrosis for BA patients during infancy.

✓ The iBALF and BALF scores can monitor liver fibrosis in a similar manner before and after 1 year of age, respectively.

✓ The BA patients with an iBALF score > 6 at presentation had poor outcome on native liver survival at 1 year of age.

1. Hartley JL, Davenport M, Kelly DA. Biliary atresia. *Lancet* 2009; **374**: 1704–1713.
2. Haafiz AB. Liver fibrosis in biliary atresia. *Expert Rev Gastroenterol Hepatol* 2010; **4**: 335–343.
3. Lykavieris P, Chardot C, Sokhn M *et al.* Outcome in adulthood of biliary atresia: a study of 63 patients who survived for over 20 years with their native liver. *Hepatology* 2005; **41**: 366–371.
4. Tomita H, Masugi Y, Hoshino K *et al.* Long-term native liver fibrosis in biliary atresia: development of a novel scoring system using histology and standard liver tests. *J Hepatol* 2014; **60**: 1242–1248.
5. Bedossa P, Poynard T. An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group. *Hepatology* 1996; **24**: 289–293.
6. Ichida F, Tsuji T, Omata M *et al.* New Inuyama classification; new criteria for histological assessment of chronic hepatitis. *Int Hepatol Commun* 1996; **6**: 112–119.
7. Nio M, Ohi R. Biliary atresia. *Semin Pediatr Surg* 2000; **9**: 177–186.

8. Wai CT, Greenson JK, Fontana RJ *et al.* A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. *Hepatology* 2003; **38**: 518–526.

9. Tanaka T, Yamashita A, Ichihara K. Reference intervals of clinical tests in children determined by a latent reference value extraction method. *J Jpn Pediatr Soc* 2008; **112**: 1117–1132.

10. Alexopoulos SP, Merrill M, Kin C *et al.* The impact of hepatic portoenterostomy on liver transplantation for the treatment of biliary atresia: early failure adversely affects outcome. *Pediatr Transplant* 2012; **16**: 373–378.

11. Neto JS, Feier FH, Bierrenbach AL *et al.* Impact of Kasai portoenterostomy on liver transplantation outcomes: A retrospective cohort study of 347 children with biliary atresia. *Liver Transpl* 2015; **21**: 922–927.

12. Kim SY, Seok JY, Han SJ *et al.* Assessment of liver fibrosis and cirrhosis by aspartate aminotransferase-to-platelet ratio index in children with biliary atresia. *J Pediatr Gastroenterol Nutr* 2010; **51**: 198–202.

13. Lind RC, Verkade HJ, Porte RJ *et al.* Aspartate transaminase-to-platelet ratio index is not correlated with severity of fibrosis or survival in children with biliary atresia. *J Pediatr Gastroenterol Nutr* 2012; **54**: 698.

14. Lampela H, Kosola S, Heikkila P *et al.* Native liver histology after successful portoenterostomy in biliary atresia. *J Clin Gastroenterol* 2014; **48**: 721–728.

15. Shin NY, Kim MJ, Lee MJ *et al.* Transient elastography and sonography for prediction of liver fibrosis in infants with biliary atresia. *J Ultrasound Med* 2014; **33**: 853–864.

16. Chang HK, Park YJ, Koh H *et al.* Hepatic fibrosis scan for liver stiffness score measurement: a useful preendoscopic screening test for the detection of varices in postoperative patients with biliary atresia. *J Pediatr Gastroenterol Nutr* 2009; **49**: 323–328.

17. Chongsrisawat V, Vejapipat P, Siripon N *et al.* Transient elastography for predicting esophageal/gastric varices in children with biliary atresia. *BMC Gastroenterol* 2011; **11**: 41.

18. Colecchia A, Di Biase AR, Scaioli E *et al.* Non-invasive methods can predict oesophageal varices in patients with biliary atresia after a Kasai procedure. *Dig Liver Dis* 2011; **43**: 659–663.

19. Tanabe M, Kawachi S, Obara H *et al.* Current progress in ABO-incompatible liver transplantation. *Eur J Clin Invest* 2010; **40**: 943–949.

20. Bedossa P, Dargere D, Paradis V. Sampling variability of liver fibrosis in chronic hepatitis C. *Hepatology* 2003; **38**: 1449–1457.

21. Bedossa P. The French METAVIR Cooperative Study Group. Intraobserver and interobserver variations in liver biopsy interpretation in patients with chronic hepatitis C. *Hepatology* 1994; **20**: 15–20.