

LIVER

Proposal of a new prognostic model for hepatocellular carcinoma: an analysis of 403 patients

R Tateishi, H Yoshida, S Shiina, H Imamura, K Hasegawa, T Teratani, S Obi, S Sato, Y Koike, T Fujishima, M Makuuchi, M Omata

Gut 2005;54:419–425. doi: 10.1136/gut.2003.035055

See end of article for authors' affiliations

Correspondence to:
Dr H Yoshida, Department of Gastroenterology, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan; yoshida-2im@h.u-tokyo.ac.jp

Revised version received 17 June 2004
Accepted for publication 19 June 2004

Background: The prognosis of hepatocellular carcinoma (HCC) is highly dependent on tumour extension and liver function. Recently, two new prognostic scoring systems—the CLIP score, developed by Italian investigators and the BCLC score, developed in Barcelona—have been widely used to assess prognosis in patients presenting with hepatocellular carcinoma. Each system has its own relative limitations.

Aims: To create a new prognostic scoring system which is simple, easy to calculate, and suitable for estimating prognosis during radical treatment of early HCC.

Methods: A total of 403 consecutive patients with HCC treated by percutaneous ablation at the Department of Gastroenterology, University of Tokyo Hospital, between 1990 and 1997 were used as the training sample to identify prognostic factors for our patients and used to develop the Tokyo score. As a testing sample, 203 independent patients who underwent hepatectomy at the Department of Hepato-Biliary-Pancreatic Surgery were studied. Prognostic factors were analysed by univariate and multivariate Cox proportional hazard regression.

Results: The Tokyo score consists of four factors: serum albumin, bilirubin, and size and number of tumours. Five year survival was 78.7%, 62.1%, 40.0%, 27.7%, and 14.3% for Tokyo scores 0, 1, 2, 3, and 4–6, respectively. The discriminatory ability of the Tokyo score was internally validated by bootstrap methods. The Tokyo score, CLIP score, and BCLC staging were compared by Akaike information criterion and Harrell's c index among training and testing samples. In the testing sample, the predictive ability of the Tokyo score was equal to CLIP and better than BCLC staging.

Conclusions: The Tokyo score is a simple system which provides good prediction of prognosis for Japanese patients with HCC requiring radical therapy.

The prognosis of hepatocellular carcinoma (HCC) depends on tumour extension as well as liver function. Worldwide, most patients with hepatocellular carcinoma have cirrhosis caused by chronic viral hepatitis (hepatitis C (HCV), hepatitis B virus).¹ Assessment of tumour related factors in isolation, such as the tumour node metastasis (TNM) staging,² does not accurately predict the prognosis of patients who have HCC and cirrhosis.³ The Child-Pugh classification has been widely used to evaluate liver function in cirrhotic patients, and has a relatively good correlation with prognosis,⁴ but cannot be used to predict survival in patients with HCC.

Okuda staging of HCC, established in 1985, is based on tumour size and liver function, as assessed by three of the four factors used in the Child Pugh score—namely, serum albumin, bilirubin, and the presence of ascites—and for some time has been used as the gold standard for prognostic assessment of HCC patients.⁵ However, this prognostic system was established by analysing patients mostly at an advanced stage of HCC, with a median survival of 4.1 months. With current advances in clinical practice, survival of HCC patients is now much longer, and the Okuda staging is unable to accurately predict prognosis in these patients.

Recently, several groups from Italy,⁶ Spain,⁷ France,⁸ Austria,⁹ and China¹⁰ proposed new prognostic systems for HCC. While the latter three were developed using a sample of patients with advanced disease (median survival 4–8 months), the CLIP and BCLC systems developed scores based on patients with early disease. The CLIP (Cancer of the Liver Italian Program) score showed a good correlation with

the prognosis of HCC patients receiving various treatments, including surgery, percutaneous ablation, transarterial chemo-embolisation, and liver transplantation.^{6–12} Along with liver function, as assessed by Child-Pugh stage, three tumour factors were included. Tumour morphology was divided into three categories: a single tumour $\leq 50\%$ of the size of the liver (score 0); multinodular HCC $\leq 50\%$ of the size of the liver (score 1); and massive HCC or tumour $>50\%$ of the size of the liver (score 2). As in Okuda staging, tumour size was broadly divided at 50% of liver volume. Seropositivity for α fetoprotein (AFP) and the presence of portal vein thrombosis were also included. However, due to advances in liver imaging techniques, especially ultrasound and computed tomography, HCC can now be detected at a much smaller size, usually smaller than 5 cm in diameter, and tumours smaller than 2 cm are frequently diagnosed. Tumour size is associated with the pathological grade of HCC, the probability of vascular invasion, and also with the prognosis of HCC patients after potentially curative treatments such as surgical resection and medical ablation.^{13–14} However, it is not known whether a HCC of 2 cm is a determinant of prognosis as previous models have not discriminated between large and small tumours.

Abbreviations: HCC, hepatocellular carcinoma; TNM, tumour node metastasis; CLIP, Cancer of the Liver Italian Program; BCLC, Barcelona Clinic Liver Cancer; AIC, Akaike information criterion; PEIT, percutaneous ethanol injection therapy; PMCT, percutaneous microwave coagulation therapy; TAE, transcatheter arterial embolisation; AST, aspartate aminotransferase; AFP, α fetoprotein; HCV, hepatitis C virus; HVPG, hepatic venous pressure gradient

Table 1 Definition of the Okuda staging system for hepatocellular carcinoma

	Points	
	0	1
Tumour size	<50% of liver	>50% of liver
Ascites	No	Yes
Albumin (g/dl)	≥3	<3
Bilirubin (mg/dl)	<3	≥3

Okuda stage I, 0 points; Okuda stage II, 1 or 2 points; Okuda stage III, 3 or 4 points.

More recently, another staging system, the BCLC (Barcelona Clinic Liver Cancer), was developed and based on both advanced and early HCC, dividing HCC into four early stages (A1–A4) and three more advanced ones (B–D). It contains elements of both the Okuda and Child-Pugh classifications. Subclassification of early stages requires formal measurement of hepatic venous pressure gradient (HVPG), which is not applicable in all patients, although clinical parameters (splenomegaly, etc) are now frequently applied.⁷ In the present study, we sought to establish a new scoring system that provides a more precise prediction of prognosis in patients with early stage HCC.

PATIENTS AND METHODS

Training sample

Between January 1990 and December 1997, 403 patients with naïve HCC received medical ablation, either percutaneous ethanol injection therapy (PEIT) or percutaneous microwave coagulation therapy (PMCT), at the Department of Gastroenterology, University of Tokyo Hospital. Their prognosis was followed up until August 2001, and survival data were used as the training samples in this study. HCC was detected with ultrasound and/or computed tomography, and

confirmed histopathologically by percutaneous tumour biopsy. Inclusion criteria for ablation were as follows: total bilirubin <3 mg/dl; platelet count >4×10³/mm³; prothrombin activity >35%; and no intractable ascites. Although most investigators performed PEIT in early stage HCC, such as for a single nodule of ≤5 cm in diameter or less than three nodules ≤3 cm in diameter,^{15 16} we did not limit the indication for ablation to tumour size alone. Patients received ablation therapy because surgery was not an option in terms of impairment in liver function, or they voluntarily chose ablation after informed consent although surgery was also possible. The ablation procedures have been described previously.^{17 18}

The following variables obtained at initial ablation therapy were used: age; sex; treatment modality (PEIT, PMCT, with or without transcatheter arterial embolisation (TAE)); tumour factors, including size, number of nodules, lobar distribution, and presence of extrahepatic metastasis; clinical manifestations, including ascites and hepatic encephalopathy; laboratory data, including albumin, bilirubin, prothrombin activity, aspartate aminotransferase (AST), alanine aminotransferase, platelet count, and AFP; positivity for viral markers (hepatitis B surface antigen and anti-hepatitis C antibody); and alcohol consumption. Okuda stage, CLIP score, and BCLC staging were also calculated using these variables (tables 1–3). We substituted the presence of either oesophageal varices or splenomegaly with platelet count less than 100 000/mm³ for HVPG ≥10 mm Hg, as described by Llovet and colleagues.⁷

All patients were placed under strict observation for recurrence of HCC, with regularly repeated ultrasound, computed tomography, and determination of serum tumour markers. If recurrence of HCC was detected, patients received additional treatments whenever possible—tumour ablation, TAE, or systemic chemotherapy. For survival analysis, the end point was death, and survival was censored on 31 August 2001.

Table 2 Definition of the Cancer of the Liver Italian Program (CLIP) scoring system for hepatocellular carcinoma

	Score		
	0	1	2
Child-Pugh stage	A	B	C
Tumour morphology	Uninodular and extension ≤50%	Multinodular and extension ≤50%	Massive or extension >50%
AFP (ng/ml)	<400	≥400	
Portal vein thrombosis	No	Yes	

AFP, α fetoprotein.

Table 3 Definition of the Barcelona Clinic Liver Cancer (BCLC) staging for hepatocellular carcinoma

BCLC stage	PST	Tumour status		
		Tumour stage	Okuda stage	Liver function status
Stage A: early HCC	0			
A1	0	Single	I	No portal hypertension and normal bilirubin
A2	0	Single	I	Portal hypertension and normal bilirubin
A3	0	Single	I	Portal hypertension and abnormal bilirubin
A4	0	3 tumours <3 cm	I–II	Child-Pugh A–B
Stage B: intermediate HCC	0	Large multinodular	I–II	Child-Pugh A–B
Stage C: advanced HCC	1–2*	Vascular invasion or extrahepatic spread*	I–II	Child-Pugh A–B
Stage D: end stage HCC	3–4†	Any		Child-Pugh C

Stages A and B: all criteria should be fulfilled.
 Stage C: at least one criterion; *PST 1–2 or vascular invasion/extrahepatic spread.
 Stage D: at least one criterion; †PST 3–4 or Okuda stage III/Child-Pugh C.

Testing sample

Between October 1994 and December 1999, 203 patients with naïve HCC underwent hepatectomy at the Department of Hepato-Biliary-Pancreatic Surgery, University of Tokyo Hospital, and their survival data served as the testing sample. Indications for hepatectomy and selection of the area for resection were previously published.¹⁹ Briefly, the surgical procedure was determined according to residual liver function, as determined by the severity of ascites, serum level of bilirubin, and indocyanine green retention rate at 15 minutes. Patients with a bilirubin level >2 mg/dl or with intractable ascites were contraindicated for hepatectomy. Patients were followed up as described above, and survival was censored on 31 August 2001.

Establishing a new prognostic score

We sought to construct a new prognostic model based on the following principles.

- (1) It is preferable to have two break points for continuous variables such as tumour size or serum albumin concentration because their distribution is wide and a single break point may not be optimal.
- (2) Variables must be those commonly assessed in clinical practice to enable comparison between different institutions.
- (3) The model should not include established classifications because they may be modified in the future, as was the case for TNM staging, and different versions may be confused.

We used survival time as the only end point in this analysis. Firstly, we performed univariate Cox proportional hazard regression to assess the statistical significance of each candidate potential factor, and the factor was retained if a significance level of $p < 0.05$ was attained. Polychotomous categorical data were represented by corresponding binary dummy variables. Continuous variables, such as serum concentration of albumin and size (diameter) of the tumour, were transformed into categorical variables. We divided each of these continuous variables into two or three levelled categorical data by setting one or two break point(s), respectively, which were then represented by one or two binary variable(s); p values were calculated for each set of break points with univariate or multivariate Cox proportional hazard regression, and the set of break points showing the lowest p value was retained if the value reached significance.

Factors showing statistical significance as a predictor were further analysed using a multivariate Cox proportional hazard regression model with stepwise selection of variables based on the Akaike information criterion (AIC). AIC is a measure of the goodness of fit (log likelihood) with a “penalty score” for the complexity of the model (number of variables included), defined as

$$\text{AIC} = -2 \times (\text{maximum log likelihood}) + 2 \times (\text{total number of parameters}),$$

and the optimum (that is, simplest effective) model gives the lowest AIC value.²⁰

A new prognostic score, designated the Tokyo score, was established, assigning ordinal scores (0, 1 and 2) to each of the selected factors according to the estimated regression coefficient in the final model.

Internal validation

We used the bootstrap method for internal validation of the Tokyo score system.²¹ Bootstrap validation is a method of random re-sampling from a given set of samples to simulate the effect of drawing samples from the same population. A re-sampled data set of the same size as the original (training)

data set was obtained by random sampling with replacement—in other words, each sample can be drawn more than once or not at all. Differences in three and five year survival rates were calculated between each pair of contiguous stages (for example, between Tokyo scores 1 and 2) using Kaplan-Meier estimation. Mean and 95% confidential interval of the difference in three and five year survival rates between the stages were determined by 2000 times iteration of such re-sampling.

External validation

We validated the Tokyo score in the testing sample as well as in the training sample with AIC and Harrell’s c index.²² Firstly, AIC was calculated in a Cox proportional regression model containing Tokyo, CLIP, and BCLC stages. Then, AIC was recalculated after removing each one of the scores, and the changes in AIC were compared. The c index is equivalent to the area under the receiver operator characteristic curve, and ranges from 0.0 to 1.0. A c index of 1 indicates perfect

Table 4 Baseline characteristics of the training sample (n = 403)

Variable	n (%)
Age (y) (median (range))	64 (35–87)
Male sex	293 (72.7%)
Viral infection	
HBsAg positive	39 (9.6%)
Anti-HCVAb positive	336 (83.4%)
Both positive	6 (1.5%)
Both negative	34 (8.4%)
Child-Pugh classification	
Class A	222 (55.1%)
Class B	160 (39.7%)
Class C	21 (5.2%)
Tumour characteristics	
Size of tumour (cm)	
<2	134 (33.3%)
2–5	234 (58.1%)
>5	35 (8.7%)
No of nodules	
Single	215 (53.3%)
2–3	135 (33.5%)
>3	53 (13.2%)
Portal invasion, present	4 (1.0%)
AFP (ng/ml)	
<100	308 (76.4%)
100–400	58 (14.3%)
>400	37 (9.2%)
Okuda stage	
I	296 (73.4%)
II	107 (26.6%)
III	0 (0%)
CLIP score	
0	118 (29.3%)
1	165 (40.9%)
2	98 (24.3%)
3	18 (4.5%)
4 ≤	4 (1.0%)
BCLC staging	
A1	70 (17.4%)
A2	58 (14.4%)
A3	32 (7.9%)
A4	100 (24.8%)
B	115 (28.5%)
C	7 (1.7%)
D	21 (5.2%)
Treatment modalities	
PEIT	363 (90.0%)
PMCT	26 (6.5%)
PEIT + PMCT	14 (3.5%)
Combined with TAE	54 (13.4%)

CLIP, Cancer of the Liver Italian Program; BCLC, Barcelona Clinic Liver Cancer; PEIT, percutaneous ethanol injection therapy; PMCT, percutaneous microwave coagulation therapy; TAE, transcatheter arterial embolisation; AFP, α fetoprotein; HBsAg, hepatitis B surface antigen; HCVAb, anti-hepatitis C virus antibody.

Table 5 Univariate analysis

Variable (n)	Hazard ratio (95% CI)	p Value
Age (y) ≥65	1.34 (1.04–1.73)	0.026
Male sex (293)	1.02 (0.769–1.36)	0.87
HBsAg positive (39)	1.04 (0.685–1.57)	0.87
Anti-HCVAb positive (336)	0.682 (0.494–0.943)	0.02
Drinking >80 g/day (90)	1.09 (0.807–1.47)	0.58
Ascites present (86)*	2.39 (1.82–3.15)	<0.0001
Encephalopathy present (27)	1.42 (1.14–1.76)	0.0015
Albumin (g/dl)		
>3.5 (192)	1	
2.8–3.5 (183)	1.99 (1.52–2.59)	<0.0001
<2.8 (28)	3.13 (2.01–4.88)	<0.0001
Bilirubin (mg/dl)		
<1 (271)	1	
1–2 (110)	1.19 (1.03–1.36)	0.016
>2 (22)	1.48 (1.26–1.72)	<0.0001
Prothrombin time		
>70% (225)	1	
40–70% (172)	1.36 (0.1.05–1.77)	0.021
<40% (6)	4.43 (1.79–10.99)	0.0013
Platelet count ≥10×10 ³ /mm ³ (173)	0.640 (0.494–0.829)	0.00063
AST >80 IU/l (178)	0.798 (0.620–1.03)	0.079
ALT >80 IU/l (166)	0.771 (0.598–0.994)	0.045
Size of tumour (cm)		
<2.0 (134)	1	
2.0–5.0 (234)	2.31 (1.70–3.13)	<0.0001
>5.0 (35)	3.73 (2.37–5.86)	<0.0001
No of nodules		
Single (215)	1	
2–3 (135)	1.18 (0.896–1.57)	0.24
>3 (53)	2.13 (1.499–3.03)	<0.0001
Lobar distribution		
Unilobar (292)	1	
Bilobar (111)	1.27 (1.11–1.45)	<0.0001
Extrahepatic metastasis present (3)	20.1 (6.12–66.3)	<0.0001
Portal vein invasion present (4)	14.6 (4.5–47.6)	<0.0001
AFP (ng/ml)		
<100 (308)	1	
100–400 (58)	1.99 (1.43–2.76)	<0.0001
>400 (37)	3.57 (2.46–5.18)	<0.0001

*Diuretic controllable ascites were included.
 AST, aspartate aminotransferase; ALT, alanine aminotransferase; AFP, α fetoprotein; HBsAg, hepatitis B surface antigen; HCVAb, anti-hepatitis C virus antibody.

concordance between the two variables (that is, the order of survival time and magnitude of prognostic score in the current study) while an index of 0.5 indicates a chance association.

Statistics

Data are expressed as mean (SD) unless otherwise specified. All statistical analyses were performed with S-plus 2000

Table 6 Multivariate analysis

Variable	Hazard ratio (95%CI)	p Value
Albumin (g/dl)		
>3.5	1	
2.8–3.5	1.74 (1.31–2.30)	0.00014
<2.8	2.45 (1.55–3.88)	0.00013
Bilirubin (mg/dl)		
<1	1	
1–2	1.40 (1.06–1.85)	0.02
>2	2.40 (1.47–3.92)	0.00049
Size of tumour (cm)		
<2.0	1	
2.0–5.0	2.21 (1.63–3.01)	<0.0001
>5.0	3.46 (2.20–5.45)	<0.0001
No of nodules		
≤3	1	
>3	1.88 (1.34–2.64)	<0.001

Table 7 Tokyo score

	Score		
	0	1	2
Albumin (g/dl)	>3.5	2.8–3.5	<2.8
Bilirubin (mg/dl)	<1	1–2	>2
Tumour size (cm)	<2	2–5	>5
Tumour No	≤3		>3

(MathSoft Inc., Seattle, Washington, USA). Statistical significance was set at p<0.05.

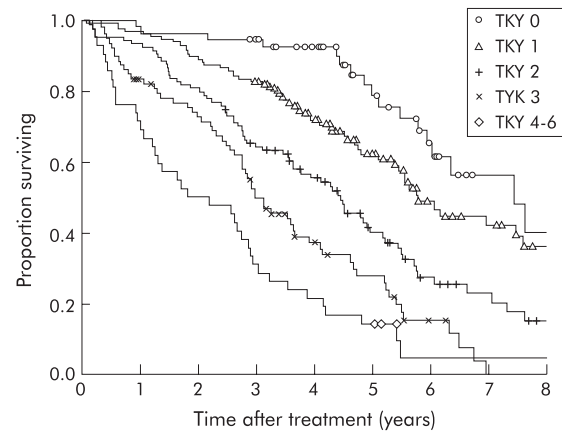
RESULTS

Patient profiles in the training sample

The training sample contained data from 293 male and 110 female patients. Baseline characteristics of the patients are shown in table 4. Median age was 64 years, with 25% and 75% percentiles at 59 and 69 years, respectively. The majority (83.4%) were HCV positive. The observation period was 3.9 (2.1) years, during which 250 patients died. Estimated 50% survival time was 4.75 years. Only eight patients (2%) were lost to follow up.

Selection of predictive factors

Univariate Cox proportional hazard analysis of the training data set revealed that 15 factors were significantly associated with prognosis of HCC patients (table 5). We included AST, which showed marginal significance (p = 0.079), and performed multivariate analysis on a total of 16 factors with stepwise selection of variables using the AIC. There were four variables which retained significance as independent predictors—namely, serum concentration of albumin, as ranked by 3.5 g/dl and 2.8 g/dl, bilirubin concentration (1 mg/dl and 2 mg/dl), size of the tumour (diameters of 2 cm and 5 cm), and number of tumour nodules (1–3 v >3) (table 6). Scores were assigned to each of the four factors according to the estimated regression coefficient in the final model (table 7) and the Tokyo score was defined as the sum of each score.



No at risk	0	1	2	3	4	5	6	7	8
TKY 0	53	53	52	51	47	28	19	12	6
TKY 1	126	123	114	105	73	47	27	18	13
TKY 2	104	98	85	66	49	29	17	10	7
TKY 3	78	66	55	37	23	15	8	0	
TKY 4-6	42	31	22	14	10	7	2	2	2

Figure 1 Kaplan-Meier estimated survival curves by Tokyo score (TKY).

Table 8 Pairwise comparisons of three and five year survival rates (with 95% confidence interval (CI)) between each stage of the Tokyo score

Stage	3 year survival (%)		5 year survival (%)	
	Difference	95% CI	Difference	95% CI
Tokyo 0 and 1	11.6	2.02–20.3	16.4	3.80–32.1
Tokyo 1 and 2	18.4	6.86–29.5	22.0	8.75–36.8
Tokyo 2 and 3	14.7	0.204–28.7	12.5	2.53–26.8
Tokyo 3 and 4–5	18.8	0.08–35.9	13.2	2.38–28.3

Internal validation

Among the training sample, 55, 126, 104, 78, 30, 9, and 3 patients were classified as Tokyo scores 0, 1, 2, 3, 4, 5, and 6, respectively. Observed cumulative survival of patients grouped by Tokyo score was calculated using the Kaplan-Meier method (fig 1). Prognosis was well distributed among the groups based on the Tokyo score. Five year survival rates for Tokyo scores 0, 1, 2, 3, and 4–6 were 78.7%, 62.1%, 40.0%, 27.7%, and 14.3%, respectively. This was confirmed by internal validation where differences in three and five year survival rates were calculated, along with 95% confidence interval, between each pair of two contiguous stages using the bootstrap method. The lower confidence limit of difference between each pair of two contiguous stages was greater than zero, indicating that all differences were statistically significant (table 8). The Tokyo score was therefore shown to be highly robust in estimating prognosis in distinct groups.

External validation

Baseline characteristics of the patients in the testing sample, who underwent surgical resection, are shown in table 9. Median age and sex proportions were similar to those in the training sample while these patients had better liver function reserve and the average tumour size was larger. Sixty five patients died during the observation period and median survival time was 5.7 years. Tokyo, CLIP, and BCLC stages were calculated according to the variables obtained from each patient.

The Tokyo Score, CLIP score, and BCLC staging were compared in both the training and testing samples by evaluating the AIC on Cox proportional hazard regression models. Goodness of fit of the model estimated by AIC was improved by removing BCLC from the model containing Tokyo, CLIP, and BCLC. AIC was greater in the model with either the Tokyo score or CLIP score alone than in the model containing both (table 10). These results indicate that the Tokyo and CLIP scores complement each other whereas addition of BCLC resulted in no improvement to the model. However, the increment was smaller when the CLIP score was removed, indicating that the model with the Tokyo score was more informative than that with the CLIP score. The c indices for the Tokyo score, CLIP score, and BCLC staging were 0.733, 0.707, and 0.657 in the testing sample and 0.737, 0.758, and 0.710 in the training sample, indicating that the Tokyo score was steadily effective in patients from different backgrounds.

DISCUSSION

The aim of this study was to create a novel, simple, prognostic scoring system that would provide a precise prediction of prognosis for patients who were candidates for radical therapy, such as percutaneous ablation or surgical resection. In addition, this scoring system would be used to stratify patients to enable comparison of the efficacy of distinct

Table 9 Baseline characteristics of the testing sample (n = 203)

Variable	n (%)
Age (y) (median (range))	64 (13–83)
Male sex	160 (78.8%)
Viral infection	
HBsAg positive	30 (14.8%)
Anti-HCVAb positive	138 (68.0%)
Both positive	2 (1.0%)
Both negative	37 (18.2%)
Child-Pugh classification	
Class A	155 (76.4%)
Class B	48 (23.6%)
Class C	0 (0%)
Tumour characteristics	
Size of tumour (cm)	
<2	28 (13.8%)
2–5	118 (58.1%)
>5	57 (28.1%)
No of nodules	
Single	146 (71.9%)
2–3	46 (22.7%)
>3	11 (5.4%)
Portal invasion present*	14 (6.9%)
AFP (ng/ml)	
<100	132 (65.0%)
100–400	29 (14.3%)
>400	42 (20.7%)
Okuda stage	
I	184 (90.6%)
II	19 (9.4%)
III	0 (0%)
CLIP score	
0	86 (42.4%)
1	67 (33.0%)
2	43 (21.2%)
3	4 (2.0%)
4≤	3 (1.5%)
BCLC staging	
A1	84 (41.4%)
A2	31 (15.3%)
A3	20 (9.9%)
A4	21 (10.3%)
B	33 (16.3%)
C	14 (6.9%)
D	0 (0%)
Tokyo score	
0	11 (5.4%)
1	59 (29.0%)
2	73 (36.0%)
3	45 (22.2%)
4≤	15 (7.4%)

*Based on imaging diagnosis.

CLIP, Cancer of the Liver Italian Program; BCLC, Barcelona Clinic Liver Cancer; AFP, α fetoprotein; HBsAg, hepatitis B surface antigen; HCVAb, anti-hepatitis C virus antibody.

Table 10 Comparison of the Tokyo score with the Cancer of the Liver Italian Program (CLIP) score and Barcelona Clinic Liver Cancer (BCLC) staging in the training and testing samples

Model	AIC	Δ AIC
Training sample (n = 403)		
Tokyo+CLIP+BCLC	2521.560	
Tokyo+CLIP	2519.722	–1.838
Removing CLIP	2523.301	+3.579*
Removing Tokyo	2552.765	+33.043†
Testing sample (n = 203)		
Tokyo+CLIP+BCLC	710.4624	
Tokyo+CLIP	709.4629	–0.9995
Removing CLIP	714.6591	+4.1967*
Removing Tokyo	715.2137	+4.7513†

*Difference in AIC (Akaike information criterion) between a model containing Tokyo and CLIP and that containing only Tokyo.

†Difference in AIC between a model containing Tokyo and CLIP and that containing only CLIP.

treatments or among different institutions. Hence the ideal staging system should provide maximal discrimination of outcomes between different stages of disease while keeping the variability of outcomes within each stage to a minimum. The Okuda staging system is not applicable to HCC patients at an early stage of disease as these patients are now eligible for potentially curative treatments, such as medical ablation or surgical resection. As such, while it was useful when first devised, the Okuda scoring or staging system is now generally considered obsolete. Thus 75% of the patients in the training sample and 90% in the testing sample were classified as Okuda stage I. Further stratification of the group is clearly needed. Recently proposed systems from France,⁸ Austria,⁹ and China¹⁰ were developed in patients with advanced disease with a median survival of only 4–8 months. We calculated the scores using the French, Austrian, and Chinese criteria in the training sample but all of these systems classified patients into only two stages, and most patients belonged to stage I (data not shown).

One of the outstanding merits of Okuda staging is the fact that it consists of four simple parameters—namely, tumour size, ascites, serum albumin, and bilirubin. In constructing the Tokyo score, we followed this simplistic approach. We examined only those parameters that are easily obtainable and avoided criteria that are not generally available. Moreover, we used the AIC in the selection of parameters to obtain a simple model without too many independent parameters or complicated computation. We believe this quality to be very important, especially in projecting retrospective analysis without omitting any patient because of missing values.

We adopted two tumour factors, size and number, in the final model. Tumour size was divided into three groups with break points of 2 cm and 5 cm in diameter. Several studies have identified tumour size in this range as significant,^{7 13 14 23 24} and a strong correlation with microvascular invasion and pathological grade of malignancy has been demonstrated.^{23 25} However, previously proposed scoring or staging systems, except for BCLC staging, did not include tumour size,⁸ or else divided the tumour broadly into massive and other, usually using the break point of half the volume of the liver.^{5 6 9} Thus differences in the present and previous scoring systems are primarily seen in the characteristics of the target populations, as reflected by our objective of defining a prognostic scoring system which would discriminate between the relatively early stages of HCC. By the same token, we did not adopt tumour related symptoms, which were included in the BCLC and French systems,^{7 8} as almost all of our patients were asymptomatic.

The number of tumour nodules is also known to be associated with intrahepatic spread of malignant cells. Some authors reported a major difference in prognosis between solitary and multinodular HCC after surgical resection.^{13 26 27} However, we found that the number of nodules was best divided dichotomously between 1–3 and ≥ 4 . The suggestion is that the presence of two or three nodules might often be the result of simultaneous independent carcinogenesis rather than intrahepatic metastasis in patients with advanced cirrhosis.

Evident vascular invasion such as portal vein tumour thrombus is an absolute predictor of ominous prognosis.^{3 13 26} Overt metastases to extrahepatic organs or lymph nodes are also associated with poor prognosis. As the testing sample contained few patients with these two manifestations, they were not selected after stepwise variable selection. It should be noted that the Tokyo score may not be predictive for advanced disease.

We selected two factors, albumin and bilirubin, as indicators of liver function. Both are included in the Child-Pugh

classification, together with prothrombin time, ascites, and encephalopathy, and thus were also included in the CLIP score as a factor in the Child-Pugh classification. However, the latter three were not selected after stepwise variable selection because they were strongly correlated with the former two. Thus liver function is represented by two parameters, which we believe is preferable for the sake of model simplicity.

Portal hypertension is accepted as a strong predictor of poor prognosis. Among our candidate factors, ascites, encephalopathy, and platelet count were considered as related to portal hypertension and all were significant in the univariate analysis. Bruix *et al* reported that HVPG was a significant predictor of decompensation after hepatic resection²⁸ and it is included in the BCLC staging system. However, HVPG is a special examination which is not routinely carried out in daily practice. We substituted the presence of oesophageal varices and platelet count less than 100 000/mm³ for HVPG ≥ 10 mm Hg, as described by the author, but the substitution may have impaired the prognostic power of the BCLC staging in the samples.

Prognostic scores can be divided into two groups: those based on expert opinion, such as TNM staging, and those developed through regression analysis of actual data, such as the CLIP and Tokyo scores. We applied bootstrap methods to avoid possible overfitting bias that often accompanies regression analysis. Nevertheless, the fact that the Tokyo score fitted better in the training sample than in the testing sample may indicate that there remained some overfitting bias. Another possible reason why the Tokyo score did not surpass the CLIP score in the testing sample was the presence of AFP in the latter but not in the former. Over 20% of patients in the testing sample had AFP levels >400 ng/ml compared with 10% in the training sample. It is reasonable to assume that AFP plays a more important role in advanced disease.

BCLC staging, developed from several independent studies on both early and advanced patients, includes treatment strategy, indicating that a single HCC without portal hypertension should be resected and that patients with no more than three nodules not exceeding 3 cm in diameter have indications for ablation therapy. Recently, Cillo *et al* found that BCLC was the best among staging systems, including CLIP, in patients treated with radical therapies.²⁹ One possible reason why BCLC staging did not show greater ability in the testing and training samples may be that our patients were not always treated according to the strategy.

The relative prognostic importance of each factor depends on the features of the patients in the training sample, as do the independent variables remaining after stepwise selection. Child-Pugh classification and the model for end stage liver disease (MELD) are suitable for assessing the prognosis of patients with severely impaired liver function³⁰ while only tumour related factors are relevant in assessing outcome after liver transplantation for HCC patients.^{31 32} Similarly, the applicability of the Tokyo score is limited by the fact that it was established and validated on the basis of HCC patients treated by medical ablation or surgical resection. However, with the growing realisation of high risk groups for HCC and rapid advances in imaging techniques, an increasing number of patients are being diagnosed at an earlier stage, and qualify for potentially curative treatments, such as medical ablation and surgical resection.

In conclusion, we established the Tokyo score by analysing survival time among HCC patients treated with medical ablation, and validated it in patients who underwent surgical resection. The Tokyo score may be useful in predicting the prognosis of HCC patients who are candidates for these curative treatments.

.....

Authors' affiliations

R Tateishi, H Yoshida, S Shiina, T Teratani, S Obi, S Sato, Y Koike, T Fujishima, M Omata, Department of Gastroenterology, University of Tokyo, Tokyo, Japan

H Imamura, K Hasegawa, M Makuuchi, Department of Hepato-Biliary-Pancreatic Surgery, University of Tokyo, Tokyo, Japan

Conflict of interest: None declared.

REFERENCES

- 1 **Shiratori Y, Shiina S, Imamura M, et al.** Characteristic difference of hepatocellular carcinoma between hepatitis B- and C- viral infection in Japan. *Hepatology* 1995;**22**:1027–33.
- 2 **Sobin L, Wittekind C.** *TNM classification of malignant tumours*, 5th edn., New York: John Wiley & Sons 1997.
- 3 **Poon RT, Fan ST, Ng IO, et al.** Prognosis after hepatic resection for stage IVA hepatocellular carcinoma: a need for reclassification. *Ann Surg* 2003;**237**:376–83.
- 4 **Pugh RN, Murray-Lyon IM, Dawson JL, et al.** Transection of the oesophagus for bleeding oesophageal varices. *Br J Surg* 1973;**60**:646–9.
- 5 **Okuda K, Ohtsuki T, Obata H, et al.** Natural history of hepatocellular carcinoma and prognosis in relation to treatment. Study of 850 patients. *Cancer* 1985;**56**:918–28.
- 6 **The Cancer of the Liver Italian Program (CLIP) Investigators.** A new prognostic system for hepatocellular carcinoma: a retrospective study of 435 patients: the Cancer of the Liver Italian Program (CLIP) investigators. *Hepatology* 1998;**28**:751–5.
- 7 **Llovet JM, Bru C, Bruix J.** Prognosis of hepatocellular carcinoma: the BCLC staging classification. *Semin Liver Dis* 1999;**19**:329–38.
- 8 **Chevre S, Trinchet JC, Mathieu D, et al.** A new prognostic classification for predicting survival in patients with hepatocellular carcinoma. Groupe d'Etude et de Traitement du Carcinome Hepatocellulaire. *J Hepatol* 1999;**31**:133–41.
- 9 **Schoniger-Hekele M, Muller C, Kutilek M, et al.** Hepatocellular carcinoma in Central Europe: prognostic features and survival. *Gut* 2001;**48**:103–9.
- 10 **Leung TW, Tang AM, Zee B, et al.** Construction of the Chinese University Prognostic Index for hepatocellular carcinoma and comparison with the TNM staging system, the Okuda staging system, and the Cancer of the Liver Italian Program staging system: a study based on 926 patients. *Cancer* 2002;**94**:1760–9.
- 11 **The Cancer of the Liver Italian Program (CLIP) Investigators.** Prospective validation of the CLIP score: a new prognostic system for patients with cirrhosis and hepatocellular carcinoma. The Cancer of the Liver Italian Program (CLIP) Investigators. *Hepatology* 2000;**31**:840–5.
- 12 **Ueno S, Tanabe G, Sako K, et al.** Discrimination value of the new western prognostic system (CLIP score) for hepatocellular carcinoma in 662 Japanese patients. Cancer of the Liver Italian Program. *Hepatology* 2001;**34**:529–34.
- 13 **The Liver Cancer Study Group of Japan.** Predictive factors for long term prognosis after partial hepatectomy for patients with hepatocellular carcinoma in Japan. The Liver Cancer Study Group of Japan. *Cancer* 1994;**74**:2772–80.
- 14 **Lencioni R, Bartolozzi C, Caramella D, et al.** Treatment of small hepatocellular carcinoma with percutaneous ethanol injection. Analysis of prognostic factors in 105 Western patients. *Cancer* 1995;**76**:1737–46.
- 15 **Ebara M, Ohto M, Sugiura N, et al.** Percutaneous ethanol injection for the treatment of small hepatocellular carcinoma. Study of 95 patients. *J Gastroenterol Hepatol* 1990;**5**:616–26.
- 16 **Livraghi T, Bolondi L, Lazzaroni S, et al.** Percutaneous ethanol injection in the treatment of hepatocellular carcinoma in cirrhosis. A study on 207 patients. *Cancer* 1992;**69**:925–9.
- 17 **Shiina S, Tagawa K, Niwa Y, et al.** Percutaneous ethanol injection therapy for hepatocellular carcinoma: results in 146 patients. *AJR Am J Roentgenol* 1993;**160**:1023–8.
- 18 **Shiina S, Teratani T, Obi S, et al.** Nonsurgical treatment of hepatocellular carcinoma: from percutaneous ethanol injection therapy and percutaneous microwave coagulation therapy to radiofrequency ablation. *Oncology* 2002;**62**(suppl 1):64–8.
- 19 **Makuuchi M, Kosuge T, Takayama T, et al.** Surgery for small liver cancers. *Semin Surg Oncol* 1993;**9**:298–304.
- 20 **Maetani S, Onodera H, Nishikawa T, et al.** Systematic computer-aided search of optimal staging system for colorectal cancer. *J Clin Epidemiol* 1991;**44**:285–91.
- 21 **Davison AC, Hinkley DV.** *Bootstrap methods and their application*. Cambridge: Cambridge University Press, 1997.
- 22 **Harrell FE Jr, Lee KL, Califf RM, et al.** Regression modelling strategies for improved prognostic prediction. *Stat Med* 1984;**3**:143–52.
- 23 **Jonas S, Bechstein WO, Steinmuller T, et al.** Vascular invasion and histopathologic grading determine outcome after liver transplantation for hepatocellular carcinoma in cirrhosis. *Hepatology* 2001;**33**:1080–6.
- 24 **Marsh JW, Dvorchik I, Bonham CA, et al.** Is the pathologic TNM staging system for patients with hepatoma predictive of outcome? *Cancer* 2000;**88**:538–43.
- 25 **Klintermalm GB.** Liver transplantation for hepatocellular carcinoma: a registry report of the impact of tumour characteristics on outcome. *Ann Surg* 1998;**228**:479–90.
- 26 **Izumi R, Shimizu K, Ii T, et al.** Prognostic factors of hepatocellular carcinoma in patients undergoing hepatic resection. *Gastroenterology* 1994;**106**:720–7.
- 27 **Nagasue N, Ono T, Yamanoi A, et al.** Prognostic factors and survival after hepatic resection for hepatocellular carcinoma without cirrhosis. *Br J Surg* 2001;**88**:515–22.
- 28 **Bruix J, Castells A, Bosch J, et al.** Surgical resection of hepatocellular carcinoma in cirrhotic patients: prognostic value of preoperative portal pressure. *Gastroenterology* 1996;**111**:1018–22.
- 29 **Cillo U, Bassanello M, Vitale A, et al.** The critical issue of hepatocellular carcinoma prognostic classification: which is the best tool available? *J Hepatol* 2004;**40**:124–31.
- 30 **Kamath PS, Wiesner RH, Malinchoc M, et al.** A model to predict survival in patients with end-stage liver disease. *Hepatology* 2001;**33**:464–70.
- 31 **Mazzafarro V, Regalia E, Doci R, et al.** Liver transplantation for the treatment of small hepatocellular carcinomas in patients with cirrhosis. *N Engl J Med* 1996;**334**:693–9.
- 32 **Yao FY, Ferrell L, Bass NM, et al.** Liver transplantation for hepatocellular carcinoma: expansion of the tumour size limits does not adversely impact survival. *Hepatology* 2001;**33**:1394–403.



Proposal of a new prognostic model for hepatocellular carcinoma: an analysis of 403 patients

R Tateishi, H Yoshida, S Shiina, H Imamura, K Hasegawa, T Teratani, S Obi, S Sato, Y Koike, T Fujishima, M Makuuchi and M Omata

Gut 2005 54: 419-425
doi: 10.1136/gut.2003.035055

Updated information and services can be found at:
<http://gut.bmj.com/content/54/3/419>

	<i>These include:</i>
References	This article cites 29 articles, 1 of which you can access for free at: http://gut.bmj.com/content/54/3/419#BIBL
Email alerting service	Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.
Topic Collections	Articles on similar topics can be found in the following collections Hepatic cancer (474)

Notes

To request permissions go to:
<http://group.bmj.com/group/rights-licensing/permissions>

To order reprints go to:
<http://journals.bmj.com/cgi/reprintform>

To subscribe to BMJ go to:
<http://group.bmj.com/subscribe/>