

**Table 3** Estimated average duration of response (days) among responders in NSCLC data

Group	Proposed method				Standard method		
	Mean	Difference	95% CI <sup>a</sup>	95% CI <sup>b</sup>	Mean	Difference	95% CI
CPT-P	89.7	16.1	-14.6, 46.7	-15.9, 48.0	86.6	9.5	-20.7, 39.7
VDS-P	73.6				77.1		

<sup>a</sup>Robust 95% confidence intervals for the difference, which account for variability of weights.

<sup>b</sup>Robust 95% confidence intervals for the difference, which do not account for variability of weights.

Our proposed approach can also estimate the size of the four potential subgroups (true responders, CPT-P only responders, VDS-P only responders and non-responders). For example, the number of true responders can be estimated by summing the estimates of weight among responders which is shown in Figure 1. The number of CPT-P only responders can be estimated by summing both the estimates of (1-weight) among responders in the CPT-P group and the estimates of weight among non-responders in the VDS-P group. Table 4 shows these results. In the NSCLC data, 13.5% of the patients were true responders, while 38.3% of the patients were non-responders who would not respond under either treatment assignment. For the CPT-P only responders, who comprised 29.7% of the population, there is a distribution of  $Y$  (duration of response) in the CPT-P assignment group. The average duration of response in the CPT-P group for patients who would not respond under the VDS-P treatment,  $E[Y_i|Z_i = 1, R_i = 1, R_i(0) = 0]$ , was estimated by weighted analysis in which weights are the probability that the response would not have occurred had the patient received the VDS-P treatment. This resulted in a point estimate of 85.1 (days) with the 95% confidence interval (69.0,101.3). Similarly, the estimated average duration of response in the VDS-P group for patients who would not respond to the CPT-P treatment,  $E[Y_i|Z_i = 0, R_i = 1, R_i(1) = 0]$ , was 79.5 (days) with the 95% confidence interval (53.8,105.3).

**Table 4** Estimated size of the four potential subgroups in NSCLC data

	VDS-P		Total
	Response	Non-response	
CPT-P	Response	73.1	106.3
		(29.7%)	
CPT-P	Non-response	94.3	139.7
		(38.3%)	
		78.6	246

### Discussion

Within the framework of Frangakis and Rubin [3], this paper develops methods for causal inference in the always-responded principal stratum, that is, we have proposed an estimation method for the “local” average treatment effect, which is the average effect of treatment among responders. Alternatively, causal inference could be made using a missing data framework that assumes all randomized subjects will respond, and thus will have the duration of response. In such an approach, the duration of response is treated as missing in patients who have not responded, and causal estimands can be defined based on the contrasts of the duration of response distributions for the standard and new treatment groups. The goal of assessing such estimands is to compare outcome distribution between the randomized groups had all patients responded during the study. Robins *et al.* [21] developed an IPCW (inverse probability of censoring weighted) method that could be used for causal inference of this estimand, which is the “global” average treatment effect in the entire study population. The drawback of such missing data approach for the present application is that the causal estimand may not be relevant or interpretable, because it is unrealistic to suppose that all patients would respond. Frangakis and Rubin [3] criticize use of such a causal estimand because it uses nonexistent “*a priori*” counterfactuals. However, there are similarities between our estimation method for outcomes in a principal stratum and the IPCW method. The underlying idea for both methods is to base estimation on the observed outcomes but weight them by some quantities for predicting each target population. The IPCW weight is the inverse of the probability of remaining in the study under the MAR (missing at random) assumption. While, under the ignorability assumption (4), our approach weights the contribution of the observed responders by their probability of response under the other treatment.

The proposed method can be used in a variety of applications where the comparison of treatments adjusted for post-treatment variables are required.

As analysed in this paper, the comparison of outcomes among patients who responded to treatment is one example. Gilbert *et al.* [6] considered a randomized study to evaluate the efficacy of a preventive HIV vaccine among infected subjects. They used a similar framework of potential outcomes to formulate the causal estimands, which were defined in terms of the distributions of potential viral loads given assignment to receive vaccine or placebo for subjects in the always-infected principal stratum. Another example is the analysis of QOL data with censoring due to death [4,5,7]. Standard approaches for missing data attempt to estimate the treatment effect that would have been observed if (contrary to fact) all patients had continued to be observed until the end of the study [21], which is the global average treatment effect. This analysis will be reasonable if the post-treatment variable is controllable. However, QOL data for censored cases due to death is "missing" because a null value exists [3]. Therefore, it is reasonable to think that estimation of the causal effects of treatment on QOL is restricted by the life of patients, which can be affected by the treatment and is a post-treatment variable.

Our estimation method is simple in that it requires only the prediction of the probability of the event in each treatment as a function of covariates. For any type of outcome variables, the proposed weighted estimator can be easily constructed among patients for whom the event has occurred and can be interpreted as the treatment effect among patients who would have had the event in either treatment group. This weighted analysis can be easily fitted in many standard statistical packages. However, the correctness of the causal inferences from our weighted analysis depends on the key assumption (4) that the value of the counterfactual indicator of response is independent of both the actual response status and the outcome variable of interest conditional on the covariates. This residual independence between the outcome variables and the counterfactual response is a non-identifiable assumption and is not testable from the observed data. If the covariates which satisfy the conditional independence assumption (4) are available and these covariates are entered as explanatory variables in the weight prediction model, the selection bias due to these unmeasured individual characteristics, that is, the departures from assumption (4), will be small. For example, when the time-dependent covariates or another outcome such as side effects are recorded and are related to response in addition to the baseline covariates, all these factors should be included in the weight prediction model. Therefore, for this issue, it is important to collect data on a sufficient number of covariates for the outcome variables to

ensure that the assumption of no unmeasured covariates will be at least approximately true. In our NSCLC data, sex, age, stage and performance status, which were most clinically important factors for tumour response, were used as the covariates to predict the probability of response in each group. The subject matter experts agreed that these covariates used in the analysis could be presumed to be most predictive of tumor response and response duration under either treatment group. Therefore, the assumption (4) seems to be reasonable in our example. Otherwise, it will be necessary to extend the proposed method to investigate the sensitivity of our inferences to the fundamental assumption of no unmeasured covariates [22]. Alternative approach to sensitivity analysis as means for summarizing uncertainty due to unmeasured factors is to compute bounds for the causal effect [7,23]. Zhang and Rubin [7] have derived such bounds for the principal causal effects (Equation 2), although the bounds are often so wide as to be useless for making decisions.

## Acknowledgements

We are grateful to the West CPT-11 Lung Cancer Study Group (Dr H Niitani, Dr M Fukuoka and other investigators) for providing us with these valuable data. We would like to thank Tosiya Sato, PhD for his helpful comments on this work. We also thank the referee, the associate editor, and the editor for their comments, which led to a much improved version of the paper. This research was supported in part by a Grant-in-Aid for Scientific Research (A) No. 16200022.

## References

1. **Rosenbaum PR.** The consequences of adjustment for a concomitant variable that has been affected by the treatment. *J R Stat Soc Series A - General* 1984; **147**: 656–66.
2. **Robins JM, Greenland S.** Identifiability and exchangeability of direct and indirect effects. *Epidemiology* 1992; **3**: 143–55.
3. **Frangakis CE, Rubin DB.** Principal stratification in causal inference. *Biometrics* 2002; **58**: 21–29.
4. **Rubin DB.** More powerful randomization-based *p*-values in double-blind trials with non-compliance. *Stat Med* 1998; **17**: 371–85.
5. **Rubin DB.** Comment on "Causal inference without counterfactuals", by A. P. Dawid. *J Am Stat Assoc* 2000; **95**: 435–37.
6. **Gilbert PB, Bosch RJ, Hudgens MG.** Sensitivity analysis for the assessment of causal vaccine effects on viral load in HIV vaccine trials. *Biometrics* 2003; **59**: 531–41.
7. **Zhang JL, Rubin DB.** Estimation of causal effects via principal stratification when some outcomes are truncated by "death". *J Educ Behave Stat* 2003; **28**: 353–68.

8. Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables (with discussion). *J Am Stat Assoc* 1996; **91**: 444–72.
9. Rubin DB. Bayesian inference for causal effects: the role of randomization. *Annals Stat* 1978; **6**: 34–58.
10. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol* 1974; **66**: 688–701.
11. Holland PW. Statistics and causal inference (with discussion). *J Am Stat Assoc* 1986; **81**: 945–70.
12. Huber PJ. The behavior of maximum likelihood estimates under nonstandard conditions. In *Proceedings of the Fifth Berkeley Symposium in Mathematical Statistics and Probability*. University of California Press, 1976: 221–33.
13. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986; **73**: 13–22.
14. Robins JM. Marginal structural models versus structural nested models as tools for causal inference. In Halloran ME, Berry D eds. *Statistical models in epidemiology: the environment and clinical trials*. Springer-Verlag, 1999: 95–134.
15. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000; **11**: 550–60.
16. Robins JM, Rotnitzky A, Zhao LP. Estimation of regression coefficients when some regressors are not always observed. *J Am Stat Assoc* 1994; **89**: 846–66.
17. Rubin DB, Thomas N. Affinely invariant matching methods with ellipsoidal distributions. *Annals Stat* 1992; **20**: 1079–93.
18. Rubin DB, Thomas N. Characterizing the effect of matching using linear propensity score methods with normal distributions. *Biometrika* 1992; **79**: 797–809.
19. Rubin DB, Thomas N. Matching using estimated propensity scores: Relating theory to practice. *Biometrics* 1996; **52**: 249–64.
20. Negoro S, Masuda N, Takada Y et al. Randomised phase III trial of irinotecan combined with cisplatin for advanced non-small cell lung cancer. *Br J Cancer* 2003; **88**: 335–41.
21. Robins JM, Rotnitzky A, Zhao LP. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *J Am Stat Assoc* 1995; **90**: 106–21.
22. Robins JM, Rotnitzky A, Scharfstein DO. Sensitivity analysis for selection bias and unmeasured confounding in missing data and causal inference models. In: Halloran ME, Berry D eds. *Statistical models in epidemiology, the environment, and clinical trials*. Springer, 1999: 1–94.
23. Balke A, Pearl J. Bounds on treatment effects from studies with imperfect compliance. *J Am Stat Assoc* 1997; **92**: 1171–78.

## Appendix

The causal parameter  $\beta = (\beta_0, \beta_1)$  of model (6) can be obtained from the solution to the following weighted estimating equations among responders  $U(\beta, \hat{\alpha}) = \sum d(Z_i; \beta) w_i(\hat{\alpha}_{1-z}) [Y_i - g(Z_i; \beta)] = \sum U_i(\beta, \hat{\alpha}) = 0$ , where  $g(Z_i; \beta) = \beta_0(1 - Z_i) + \beta_1 Z_i$ ,  $d(Z_i; \beta) = \partial g(Z_i; \beta) / \partial \beta$ , and  $\hat{\alpha} = (\hat{\alpha}_0, \hat{\alpha}_1)$  are the maximum likelihood estimators of  $\alpha_z$  in each treatment group under the prediction model for  $\Pr(R_i = 1 | X_i, \alpha_z) = w(X_i; \alpha_z)$ . Typically,  $w(\cdot)$  would be chosen to be a logistic function. Under the assumption (4) and the true model for  $w(X_i; \alpha_z)$ , there exists a unique solution  $\hat{\beta}$ , and  $U(\beta_{\text{true}}, \hat{\alpha})$  and  $(\hat{\beta} - \beta_{\text{true}})$  are asymptotically normal with a mean of zero and respective asymptotic variances  $C$  and  $(\Gamma^{-1})C(\Gamma^{-1})^T$  that can be consistently estimated by  $\hat{C}$  and  $(\hat{\Gamma}^{-1})\hat{C}(\hat{\Gamma}^{-1})^T$ , where  $\hat{\Gamma} = \partial U(\hat{\beta}, \hat{\alpha}) / \partial \beta^T$ ,  $\hat{C} = \hat{A} - \hat{B}\hat{\Omega}\hat{B}^T$ ,  $\hat{A} = \sum U_i(\hat{\beta}, \hat{\alpha}) U_i(\hat{\beta}, \hat{\alpha})^T$ ,  $\hat{B} = \partial U(\hat{\beta}, \hat{\alpha}) / \partial \alpha^T$ , and  $\hat{\Omega}$  is the estimate of  $\Omega$  based on the observed information from the likelihood for the prediction model  $w(X_i; \alpha_z)$ . The above variance estimator for  $\beta$  can be expressed as  $(\hat{\Gamma}^{-1})\hat{A}(\hat{\Gamma}^{-1})^T(\hat{\Gamma}^{-1})\hat{B}\hat{\Omega}\hat{B}^T(\hat{\Gamma}^{-1})^T$ . The first term  $(\hat{\Gamma}^{-1})\hat{A}(\hat{\Gamma}^{-1})^T$  is the usual robust (sandwich) variance estimate without accounting for the variability of weights, thus, the estimator  $\hat{\beta}$  that uses the estimated weights is at least as efficient as the one based on the usual robust variance estimate.

# Prospective Cohort Study of Transarterial Chemoembolization for Unresectable Hepatocellular Carcinoma in 8510 Patients

KENICHI TAKAYASU,\* SHIGEKI ARII,† IWAO IKAI,<sup>§</sup> MASAO OMATA,<sup>||</sup> KIWAMU OKITA,<sup>¶</sup> TAKAFUMI ICHIDA,<sup>#</sup> YUTAKA MATSUYAMA,\*\* YASUNI NAKANUMA,<sup>††</sup> MASAMICHI KOJIRO,<sup>§§</sup> MASATOSHI MAKUUCHI,<sup>|||</sup> and YOSHIO YAMAOKA<sup>¶¶</sup> for the Liver Cancer Study Group of Japan

\*Department of Diagnostic Radiology, National Cancer Center Hospital, Tokyo; †Department of Hepato-Biliary-Pancreatic Surgery, Tokyo Medical and Dental University Graduate School of Medicine, Tokyo; ‡Department of Gastroenterological Surgery, Kyoto University Graduate School of Medicine, Kyoto; §Department of Gastroenterology, Graduate School of Medicine, University of Tokyo, Tokyo; ¶Social Insurance Shimonoseki Kohsei Hospital, Yamaguchi; #Department of Gastroenterology, Juntendo University School of Medicine, Shizuoka; \*\*Department of Biostatistics, School of Health Sciences and Nursing University of Tokyo, Tokyo; ††Department of Human Pathology, Kanazawa University Graduate School of Medicine, Kanazawa; §§Department of Pathology, Kurume University School of Medicine, Kurume; |||Hepato-Biliary-Pancreatic Surgery Division, Department of Surgery, Graduate School of Medicine, University of Tokyo, Tokyo; and ¶¶Kitano Hospital, The Tazuke Kofukai Medical Research Institute, Osaka, Japan

**Background & Aims:** To elucidate the survival of the patients with unresectable hepatocellular carcinoma (HCC) who underwent transcatheter arterial lipiodol chemoembolization (TACE) and to analyze the factors affecting the survivals. **Methods:** During the last 8 years, a nationwide prospective cohort study was performed in 8510 patients with unresectable HCC who underwent TACE using emulsion of lipiodol and anticancer agents followed by gelatin sponge particles as an initial treatment. Exclusion criteria were extrahepatic metastases and/or any previous treatment prior to the present TACE. The primary end point was survival. The survival rates were calculated by the Kaplan-Meier method. The multivariate analyses for the factors affecting survival were evaluated by the Cox proportional hazard model. The mean follow-up period was 1.77 years. **Results:** For overall survival rates by TACE, median and 1-, 3-, 5-, and 7-year survivals were 34 months, 82%, 47%, 26%, and 16%, respectively. Both the degree of liver damage and the tumor-node-metastasis (TNM) system proposed by the Liver Cancer Study Group of Japan demonstrated good stratification of survivals ( $P = .0001$ ). The multivariate analyses showed significant difference in degree of liver damage ( $P = .0001$ ),  $\alpha$ -fetoprotein value ( $P = .0001$ ), maximum tumor size ( $P = .0001$ ), number of lesions ( $P = .0001$ ), and portal vein invasion ( $P = .0001$ ). The last 3 factors could be replaced by TNM stage. The TACE-related mortality rate after the initial therapy was .5%. **Conclusions:** TACE showed safe therapeutic modality with a 5-year survival of 26% for unresectable HCC patients. The degrees of liver damage, TNM stage, and  $\alpha$ -fetoprotein values were independent risk factors for patient survival.

remains poor because of the advanced stage of cancer and associated hepatic impairment at diagnosis and because of the high intrahepatic recurrence rate at 5 years of 79%–80% after hepatic surgery<sup>2,3</sup> and of 83% after percutaneous ethanol injection (PEI),<sup>4</sup> resulting from either intrahepatic metastases from the primary tumor or multicentric occurrence. With regard to treatment, surgical resection and PEI or radio-frequency ablation (RFA) are considered to be curable treatments of choice.<sup>5</sup> Whereas transcatheter arterial embolization irrespective of with and without anticancer agent and lipiodol has been controversial in the survival benefit,<sup>6–10</sup> randomized controlled trials recently elucidated that TACE showed a survival benefit compared with control.<sup>11,12</sup>

Transcatheter arterial embolization was initially used to treat HCC by Doyon et al<sup>13</sup> in 1974 and was applied to most inoperable HCC using gelatin sponge particles and anticancer agents in Japan.<sup>14</sup> In the mid-1990s, lipiodol was newly introduced to enhance mainly the therapeutic effect.<sup>15–18</sup> Transcatheter arterial lipiodol chemoembolization (TACE) is now the mainstay of treatment of choice for noncurative HCC.<sup>19–32</sup>

The latest nationwide report registered every 2 years by the Liver Cancer Study Group of Japan<sup>33</sup> showed the frequency of treatment for HCC as an initial therapy: transcatheter arterial embolization including TACE and transcatheter arterial injection of emulsion of lipiodol and anticancer agent (without gelatin sponge) accounted

*Abbreviations used in this paper:* HCC, hepatocellular carcinoma; PEI, percutaneous ethanol injection; RFA, radio-frequency ablation; TACE, transcatheter arterial lipiodol chemoembolization; TNM, tumor-node-metastasis.

© 2006 by the American Gastroenterological Association Institute  
0016-5085/06/\$32.00  
doi:10.1053/j.gastro.2006.05.021

Hepatocellular carcinoma (HCC) is the fifth most common cancer in the world and is responsible for 500,000 deaths globally every year.<sup>1</sup> The prognosis still

for 36.4% of 16,941 patients; surgical resection for 31.3%; local ablation therapy such as PEI, microwave coagulation therapy, and RFA for 26.8%; and others for 5.5%. Moreover, TACE serves for treatment of intrahepatic recurrent foci in 52% of patients who developed after previous treatments.<sup>33</sup>

To date, the long-term survival rate and analysis of predictor affecting the survival rate of TACE for more than 8000 patients with HCC have not yet been reported. Thus, we conducted a study to elucidate them using a list of patients registered nationwide.

## Materials and Methods

### Patients

During the last 8 years from January 1994 to December 2001, a total of 72,866 patients with primary liver cancer were newly registered biannually by the Liver Cancer Study Group of Japan throughout some 800 medical institutions using a registration/questionnaire sheet with more than 180 questions. HCC patients occupied 95% (69,182/72,866) of primary liver cancer. Of these, 8542 patients were assigned in the current prospective cohort study, who underwent TACE as an initial treatment for unresectable HCC and did not receive any other therapy during the first investigation period of no more than 2 years and fulfilled the following inclusion criteria. First, the hepatic lesion was clinicopathologically diagnosed as HCC. Other primary hepatic lesions such as intrahepatic cholangiocarcinoma, sarcoma, and other benign lesions were excluded. Second, there was neither extrahepatic metastasis to lymph nodes and/or other organs nor any treatment prior to the present TACE. Finally, 8510 patients were enrolled, with the exclusion of 32 patients whose replies for the questionnaire were incomplete.

The diagnosis of HCC was made mainly based on the reliable imaging modalities using ultrasonography, dynamic computed tomography (CT), magnetic resonance imaging (MRI), and angiography and/or pathologically by biopsy specimens (4.5%). Abnormal elevation of tumor markers was also referred:  $\alpha$ -fetoprotein more than 400 ng/mL (normal, <20 ng/mL) and des- $\gamma$ -carboxyl prothrombin more than 100 mAU/mL (normal, <40 mAU/mL). Typical HCC was depicted as hyperattenuation in arterial phase and hypoattenuation or washout in delayed phase (approximately 3 minutes after the contrast injection) of dynamic CT<sup>34</sup> and on dynamic MRI, as was shown a hypervascular lesion on hepatic arteriography. The extrahepatic metastases were routinely examined by ultrasonography, CT, and chest x-ray.

The distribution of background factors was tabulated in Table 1. The study population consisted of 6122 males and 2386 females with a mean age of 66.2 years (standard deviation (SD), 9.15). The number of patients in grades A, B, and C of degree of liver damage was 4008 (51%), 3053 (39%), and 766 (10%), respectively. The degree of liver damage proposed by the Liver Cancer Study Group of Japan was classified as grades

**Table 1.** Distribution of Background Factors and Results of the Univariate Analysis in 8510 Patients With Unresectable Hepatocellular Carcinoma Who Underwent Transcatheter Arterial Lipiodol Chemoembolization

Background factors	Number of patients	Proportion (%)	Survival (%)			P value
			1 y	3 y	5 y	
Age, y						
<60	1871	22	78	44	29	.07
≥60	6639	78	83	47	25	
Sex						
Male	6122	72	81	46	25	.002
Female	2386	28	84	50	28	
Degree of liver damage						
A	4008	51	87	56	33	.0001
B	3053	39	80	41	21	
C	766	10	63	23	8	
HBV and HCV						
HCV Ab positive	6063	74	84	48	25	.0001
HBs Ag positive	900	11	73	39	24	
Both positive	211	3	80	50	32	
Both negative	971	12	78	45	28	
Maximum tumor size (cm)						
≤2	1989	24	93	63	39	.0001
2.1–3	1981	24	90	52	28	
3.1–5	2318	28	83	43	23	
≥5.1	2076	25	63	30	16	
No. of lesions						
1	3648	44	87	57	33	.0001
2–3	2675	32	84	45	24	
4</=	2066	25	72	33	16	
Degree of portal vein invasion (Vp)						
Vp0	6881	88	87	51	28	.0001
Vp1	323	4	64	23	12	
Vp2	306	4	51	22	11	
≥Vp3	347	4	35	10	—	
Degree of hepatic vein invasion (Vv)						
Vv0	7246	97	85	49	27	.0001
≥Vv1	244	3	48	15	—	
Alpha-Fetoprotein (ng/mL)						
≤20	2745	34	90	59	34	.0001
21–200	2807	35	87	49	27	
201–400	591	7	82	38	19	
401–1000	603	7	72	30	19	
≥1001	1399	17	60	27	15	
Des- $\gamma$ carboxy-prothrombin (mAU/mL)						
≤100	3557	54	90	56	33	.0001
101–299	811	12	85	45	23	
300–499	339	5	82	42	20	
500–999	406	6	77	33	24	
≥1000	1478	22	64	28	14	
TNM stage						
I (T1N0M0)	927	13	96	72	47	.0001
II (T2N0M0)	2934	40	90	57	32	
III (T3N0M0)	2949	40	80	39	20	
IV A (T4N0M0)	501	7	49	16	10	

Vp, vascular invasion of the portal vein; Vv, vascular invasion of the hepatic vein.

**Table 2.** Degree of Liver Damage Classified by the Liver Cancer Study Group of Japan

Item	Degree of liver damage		
	A	B	C
Ascites	None	Controllable	Uncontrollable
Serum bilirubin (mg/dL)	<2.0	2.0–3.0	>3.0
Serum albumin (g/dL)	>3.5	3.0–3.5	<3.0
ICGR15 (%)	<15	15–40	>40
Prothrombin activity (%)	>80	50–80	<50

NOTE. If 2 or more items scoring the same grade occurred in the 2 grades, the higher grade is adopted as the degree of liver damage. ICGR15 (%), indocyanine green retention rate at 15 minutes.

A, B, and C based on the highest grade containing at least 2 findings such as ascites, serum bilirubin, serum albumin, indocyanine green retention rate at 15 minutes (ICG R15), and prothrombin activity (Table 2).<sup>35,36</sup> This classification is closely related to the Child–Pugh class.

Seventy-four percent of patients were positive for hepatitis C virus antibody, 11% were positive for hepatitis B virus surface antigen, 3% were positive for both, and 12% were negative for both. Regarding the maximum tumor size, 24% of patients each had lesion(s) no larger than 2 cm in diameter and lesion(s) measured from 2.1 to 3 cm, 28% of patients had lesion(s) from 3.1 to 5 cm, and 25% had lesion(s) larger than 5.1 cm. The number of lesions was 1 in 44% of patients, 2 or 3 in 32% of patients, and more than 4 in 25% of patients. The degree of vascular invasion to the portal vein was Vp0 (no invasion) in 88% of patients, and in the remaining 12% of patients, Vp1 (invasion to the third order or more peripheral branch) in 4%, Vp2 (the second order portal vein) in 4%, and more than Vp3 (the first portal vein or main portal trunk) in 4% of patients. Meanwhile, the degree of hepatic vein invasion was Vv0 (no invasion) in 97% of patients and more than Vv1 (any hepatic vein invasion including the main hepatic veins and/or inferior vena cava) in 3% of patients. Normal AFP value was seen in 34% of patients, 21–200 ng/mL in 35%, 201–400 ng/mL in 7%, and more than 401 ng/mL in 25%. The des- $\gamma$ -carboxyl prothrombin value of more than 101 mAU/mL was recognized in 46% of patients.

The tumor-node-metastasis (TNM) staging adopted in the present study was revised by the Liver Cancer Study Group of Japan in 2000 (Table 3)<sup>35,36</sup> and was proposed as a new concordant TNM classification of the primary liver cancer by the International Hepato-Pancreato-Biliary Association.<sup>37</sup> Namely, T category consists of 3 factors such as a single lesion, no larger than 2 cm in diameter, and no vascular or bile duct invasion. T stage is determined based on the factors that are fulfilled from zero to 3. Each stage of I, II, III, and IV-A corresponds to the T factor. Stage IV-B is excluded because of the inclusion criterion. The present study showed 13% of patients in stage I, 40% each in stages II and III, and 7% in stage IV-A.

The catheter tip was advanced at the nearest site of the feeding artery as possible. The emulsion of anticancer agent

and lipiodol followed by gelatin sponge particles was carefully injected in the x-ray monitoring. The dose of emulsion of anticancer agent and lipiodol and the pieces of embolic materials used for TACE were determined based on the tumor size and extension of the lesions. The patients were followed by dynamic CT or MRI every 3 to 4 months, and repeated TACE was determined when the local recurrence, intrahepatic metastases, and/or second primary HCC was found.

The survival rate of patients who underwent TACE was calculated from the date of diagnosis of HCC and the follow-up was ended on December 31, 2001. All patient deaths were the end point irrespective of cause of death. The mean follow-up period was 1.77 years (range, 0.003–7.99). TACE-related death was designated as death within 30 days of the initial therapy.

### Statistical Analysis

The 11 variables listed in Table 1 were analyzed via univariate analysis. The multivariate analysis was performed by the Cox proportional hazard model. The survival rates were obtained by the Kaplan–Meier method and compared by the log-rank test. All variables with a *P* value less than .05 by univariate analysis were subjected to multivariate analysis. All significance tests were 2-tailed, and a *P* value less than .05 was considered statistically significant. All statistical analyses were performed with the Statistical Analysis System (SAS) version 8.02 (SAS Inc, Cary, NC).

### Results

During an 8-year follow-up period, the total numbers of patient deaths and censored cases were 3605 and 4905, respectively. The number of patients who died and were censored annually was as follows: 1316 and 1939 during year 1, 1067 and 1258 from 1 to 2 years, 630 and 764 from 2 to 3 years, 341 and 454 from 3 to 4 years, 148 and 216 from 4 to 5 years, 72 and 145 from 5 to 6 years, 23 and 81 from 6 to 7 years, and 8 and 48 from 7 to 8 years, respectively.

**Table 3.** Definitions of TNM Stage Proposed by the Liver Cancer Study Group of Japan

T factor	
T1	Fulfilling 3 factors
T2	Fulfilling 2 factors
T3	Fulfilling 1 factor
T4	Fulfilling 0 factor
Stages	
Stage I	T1 NO M0
Stage II	T2 NO M0
Stage III	T3 NO M0
Stage IV-A	T4 NO M0, or any T N1 M0
Stage IV-B	Any T NO-1 M1

T factor consists of I, Single; II, <2 cm; III, no vascular involvement

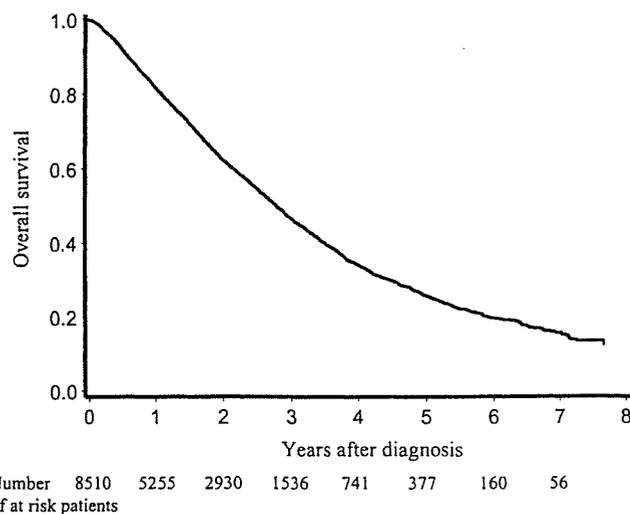
**Table 4.** The Survival of Overall Patients Who Underwent Transcatheter Arterial Lipiodol Chemoembolization, Based on Degree of Liver Damage, TNM Stage, and Combination of Degree of Liver Damage and TNM Stage

Grading/staging	n	Survival (%)							Median (mo)	P value
		1 y	2 y	3 y	4 y	5 y	7 y			
Overall survival	8510	82	63	47	34	26	16	34		
Degree of liver damage (n = 7827)										
A	4008	87	71	56	42	33	21	41	.0001	
B	3053	80	59	41	29	21	12	30		
C	766	63	37	23	15	8	—	17		
TNM stage by Liver Cancer Study Group of Japan (n = 7311)										
I	927	96	86	72	57	47	30	56	.0001	
II	2934	90	73	57	43	32	22	42		
III	2949	78	56	39	26	20	11	29		
IV-A	501	49	27	16	10	10	—	12		
Combination of degree of liver damage and TNM stage										
Liver damage A (n = 3499)										
I	489	98	92	78	64	52	38	62	.0001	
II	1439	94	80	66	52	39	26	50		
III	1358	84	64	47	33	27	14	35		
IV-A	213	59	34	24	12	10	—	15		
Liver damage B (n = 2667)										
I	309	92	82	67	50	43	16	47	.0001	
II	1068	88	70	52	38	28	16	38		
III	1116	79	54	32	22	15	8	26		
IV-A	174	45	21	13	10	10	—	11		
Liver damage C (n = 648)										
I	59	94	69	52	37	24	—	39	.0001	
II	224	72	49	29	19	11	—	24		
III	282	66	30	23	14	9	—	17		
IV-A	83	27	16	3	—	—	—	6		

**Survival Rates**

For overall survival of the 8510 patients who underwent TACE, the median and 1-, 3-, 5-, and 7-year survival rates were 34 months, 82%, 47%, 26%, and 16%, respectively (Table 4) (Figure 1). With the degree of liver damage classification, 5-year survival rates of grades A, B, and C were 33%, 21%, and 8%, respectively, with statistical significance ( $P = .0001$ ) (Figure 2). According to the TNM staging system, 5-year survival rates in stages I, II, III, and IV-A were 47%, 32%, 20%, and 10%, respectively, with significant difference among them ( $P = .0001$ ) (Figure 3).

Next, the survival rates of patients were estimated using a combination of degree of liver damage and the TNM staging system. In the degree of liver damage grade A, 5-year survival rates in stages I, II, III, and IV-A were 52%, 39%, 27%, and 10%, respectively, with significant differences among 4 subgroups ( $P = .0001$ ). The well-stratified survival curves were obtained (not shown). The same tendency as grade A was recognized in liver damage grades B and C ( $P = .0001$ ). On the other hand, 5-year survival rates in TNM stage I through liver

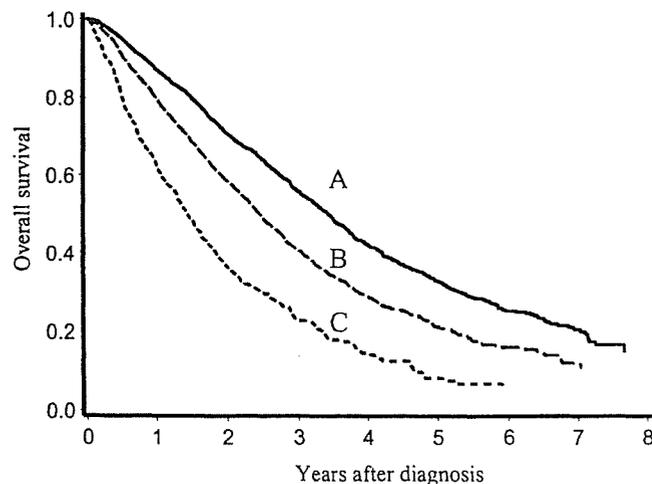


**Figure 1.** Overall survival rates of 8510 patients with unresectable HCC treated by transcatheter arterial lipiodol chemoembolization between 1994 and 2001. The 1-, 3-, and 5-year survival rates were 82%, 47%, and 26%, respectively.

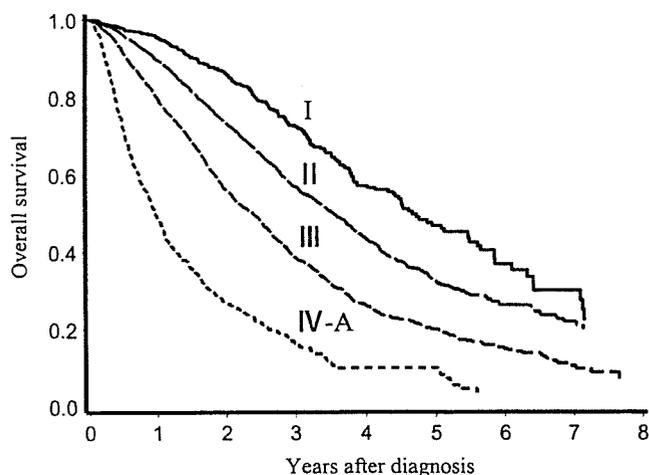
damage grades A, B, and C were 52%, 43%, and 24%, respectively, with significant differences ( $P = .0001$ ). The same results were recognized in stages II, III, and IV-A through the liver damage grades A, B, and C ( $P = .0001$ ).

**Analysis of Factors Affecting the Survival of Patients**

The univariate analysis revealed the following 11 factors as independent prognostic variables: gender, degree of liver damage, hepatitis B virus surface antigen, hepatitis C virus antibody, maximum tumor size, number of lesions, portal vein invasion, hepatic vein invasion,  $\alpha$ -fetoprotein (AFP) value, des- $\gamma$  carboxyl prothrombin



**Figure 2.** Comparison of survival rates stratified by degrees of liver damage A, B, and C with statistical significance ( $P = .0001$ ).



**Figure 3.** Comparison of survival rates stratified by TNM stages I, II, III, and IV-A with statistical significance ( $P = .0001$ ).

value, and TNM stage (Table 1). The multivariate analysis showed the following 5 factors as an independent prognostic factor: degree of liver damage, maximum tumor size, number of lesion(s), degree of portal vein invasion, and AFP value in trial 1 (Table 5). The hazard ratio of each variable ranged from 1.48 to 2.52, among which the degree of liver damage between grades A and C followed by portal vein invasion had the highest hazard ratio of 2.52 and 2.27, respectively.

Provided 3 factors of maximum tumor size, number of lesions, and portal vein invasion as seen in trial 1, were

replaced by 1 factor of TNM stage in trial 2, finally, 3 variables of degree of liver damage, TNM stage, and AFP value were statistically independent risk factors for patient survival with high hazard ratio ranging from 1.49 to 5.07 in trial 2 than from 1.49 to 2.52 in trial 1 (Table 5). TNM stage between stages I and IV-A and degree of liver damage between grades C and A had a high hazard ratio of 5.07 and 2.56, respectively. The des- $\gamma$ -carboxyl prothrombin was not evaluated because of incomplete data acquisition in the questionnaire, despite a significant factor by univariate analyses.

**TACE-Related Mortality Rate**

Forty-four (.5%) of 8510 patients developed treatment-related death after the initial TACE. The breakdown of the cause of death consisted of hepatic failure in 18 patients, cancer death in 8, intraperitoneal rupture of HCC in 7, rupture of esophago-gastric varices in 4, gastrointestinal bleeding in 1, and other causes in 6. The degree of liver damage in these 44 cases was grade A in 7 patients, grade B in 14, grade C in 17, and unknown in 6.

**Discussion**

As the background of causal factor of HCC in the current study, hepatitis C virus infection was more prevalent than hepatitis B virus (74% to 11%, respectively). The ratio of hepatitis C virus to hepatitis B virus was

**Table 5.** Multivariate Analysis of Factors Affecting Survival of Patients Who Underwent Transcatheter Arterial Lipiodol Chemoembolization Using Cox Proportional Hazard Model in Trials 1 and 2

Variables	P value	Trial 1	P value	Trial 2
		Hazard ratio (95% CI)		Hazard ratio (95% CI)
Age, y, $\geq 60$ vs $< 60$	.42	1.04 (.95-1.14)	.36	1.05 (.95-1.15)
Sex, Male vs female	.16	1.07 (.98-1.16)	.19	1.06 (.97-1.16)
Degree of liver damage				
B vs A	.0001	1.49 (1.37-1.61)	.0001	1.49 (1.37-1.62)
C vs A	.0001	2.52 (2.24-2.85)	.0001	2.56 (2.28-2.89)
HCV+ve vs HCV-ve/HBsAg-ve <sup>a</sup>	.39	.95 (.84-1.07)	.26	.93 (.82-1.05)
HBsAg+ve vs HBsAg-ve/HCV-ve <sup>b</sup>	.1	1.15 (.97-1.35)	.1	1.15 (.97-1.35)
HBsAg+ve and HCV+ve vs HBsAg-ve and HCV-ve	.84	1.03 (.79-1.34)	.99	1 (.77-1.30)
Maximum tumor size, $> 2.1$ cm vs $\leq 2$ cm	.0001	1.55 (1.40-1.71)		
No. of lesions, multiple vs solitary	.0001	1.48 (1.37-1.61)		
Portal vein invasion (Vp), present vs absent	.0001	2.27 (2.02-2.55)		
Hepatic vein invasion (Vv), present vs absent	.1	1.18 (.97-1.44)		
TNM				
II vs I			.0001	1.51 (1.29-1.77)
III vs I			.0001	2.4 (2.06-2.80)
IVA vs I			.0001	5.07 (4.21-6.10)
$\alpha$ -Fetoprotein (ng/mL), $> 401$ vs $\leq 400$	.0001	1.8 (1.65-1.96)	.0001	1.82 (1.68-1.99)

<sup>a</sup>HBsAg -ve and HCV +ve vs HBsAg -ve and HCV -ve.  
<sup>b</sup>HBsAg +ve and HCV -ve vs HBsAg -ve and HCV -ve.

almost similar to that reported by a Spanish paper<sup>11</sup> and was different from that from Hong Kong where hepatitis B virus was predominant<sup>12</sup> and from France where alcohol abuse was the main cause of HCC.<sup>8</sup> In relation to the tumor size, small HCC no larger than 2 cm in diameter was found in 24% of patients and less than 5 cm in 75% of patients. By contrast, in the countries outside Japan,<sup>11,12</sup> TACE was performed for larger HCCs—approximately 7 cm—than in those we performed. Regarding the number of lesions, solitary lesions were in 44% of patients, which was similar in incidence to that of 24%–43% reported by other studies.<sup>11,12</sup> The patients without portal vein invasion (88%) showed better survival rate than those with it, especially those with lobar or main portal tumor thrombus (portal vein trunk) of 0% at 5 years. The normal AFP value of no more than 20 ng/mL was in 34% of patients, and the abnormal elevation of more than 400 ng/mL was only in 24% of patients. These findings suggest that most HCC lesions were found as asymptomatic by periodical screening system with ultrasonography and AFP measurement for high-risk population, whereas the degree of liver function was 51% of patients in grade A and only 10% in grade C (Table 1).

The nationwide study showed that the overall survival rates of the 8510 patients who underwent TACE as an initial treatment were 47% at 3 years, 26% at 5 years, and 16% at 7 years. The previous papers reported that TACE with lipiodol showed 3-year and 5-year survival rates to be 26%–29%<sup>11,12</sup> and 16%–20%,<sup>28,31</sup> respectively, whereas other papers reported that TACE without lipiodol demonstrated various 5-year survival rates from 6% to 23%.<sup>38,39</sup> Our previous national study in 9363 patients who received TACE without lipiodol registered from 1980 to 1993 showed the 5-year survival rate to be 9%.<sup>40</sup> The marked difference of survivals between 2 different decades might depend in part on the introduction of lipiodol and the technique of super selective catheterization with the advent of the microcatheter system and in part on the lead time bias because of the advance of diagnostic modalities.

Both the degree of liver damage and the TNM staging system could stratify the survival rates with statistically significant difference ( $P = .0001$ ). The same significant differences were also recognized among TNM stages I, II, III, and IV-A in 3 different degrees of liver damage grades A, B, and C ( $P = .0001$ ). In the group with liver damage grade C, no survival remained at 7 years. By contrast, the patients with liver damage grade A and TNM stage I, the longest survival expected, demonstrated a much better 5-year survival rate of 52% than those in other subgroups. Namely, the lower the degree

of liver damage and TNM stage, the better the survival rate of patients.

Multivariate analyses revealed the following 5 variables to be independent predictor of patient prognosis such as degree of liver damage, maximum tumor size, number of lesion(s), portal vein invasion, and AFP value.<sup>41</sup> Portal vein invasion showed much higher hazard ratio than the other 4 variables. A similar result was elucidated from the study of surgical resection.<sup>42</sup> The TNM stage could replace the 3 factors of maximum tumor size, number of lesions, and portal vein invasion. The TNM staging system proposed by the Liver Cancer Study Group of Japan<sup>35,36</sup> is more accurate to classify the HCC patients in Japan than the new TNM stage proposed by the American Joint Committee on Cancer, in which a cut-off value of 5 cm is adopted as the T factor.<sup>43</sup> In general, tumor burden, hepatic functional reserve, performance status, and response to treatment have been given as prognostic factors.<sup>44–47</sup> In the current study, tumor response of TACE was not assessed because the consensus of therapeutic effectiveness of TACE with lipiodol is not still obtained.

The degree of liver damage proposed by the Liver Cancer Study Group of Japan adopts the indocyanine green 15-minute retention rate instead of hepatic encephalopathy in Child–Pugh classification and is slightly different in the cut-off value among total albumin and prothrombin activity values in 2 classifications (Table 2). The degree of liver damage could more precisely discriminate the patients with Child–Pugh class A, good candidate for surgical resection,<sup>48</sup> to grades A or B.

The policy to determine the repeated TACE seems extremely different between Japan and Western countries. We have routinely performed repeated TACE based on the findings of follow-up CT; after confirmation of whether local recurrence, intrahepatic metastases, and/or de novo evolution of HCC developed or not.<sup>39</sup> However, most randomized controlled trials designed mainly in Western countries determined repeated TACE periodically at specific intervals ranging from 2 to 3 months.<sup>8,11,12</sup> The procedure in a randomized controlled trial appears scientific to elucidate the efficacy of TACE for HCC; however, some possible drawbacks could be considered, such as the deterioration of hepatic function because of multiple procedures of TACE for a short interval and to the proximal embolization of the right, left, or proper hepatic artery.<sup>8,10,12</sup> One paper reported that hepatic failure occurred in up to 60% of patients,<sup>44</sup> whereas Ernst et al<sup>49</sup> retrospectively demonstrated that the efficacy of TACE increased when it was repeated only when necessary on follow-up CT or MRI rather than on

scheduled intervals. To our knowledge, there was no trial to assess what was better during the follow-up.

As another important factor to obtain good survival of TACE, super selective catheterization using a 2.7F to 3.0F microcatheter has been developed to maximize the therapeutic effect for targeted lesion and to minimize the unnecessary injury to the noncancerous hepatic portion.<sup>22,23,32</sup> In addition, this technique is indispensable to avoid adverse complication such as necrosis of the gallbladder or refractory gastroduodenal ulcer secondary to the embolization for nontargeted organs.<sup>29</sup> According to the latest report of the national study of the primary liver cancer,<sup>33</sup> chemoembolization was carried out by super selective catheterization: in those more peripheral to the segmental artery in 29% of 7567 patients, in 1 or 2 segmental arteries in 42%, in the right or left hepatic artery in 20%, and in the proper hepatic artery only in 9%.

TACE-related mortality rate was .5% (40/8510 patients) after the initial treatment. The previous study reported that TACE-related mortality rate was 10% (3/30 patients) in 1991,<sup>50</sup> 9.4% (12/127 patients) in 1994,<sup>26</sup> and, recently, 1.1% (2/182 patients) in 1999,<sup>31</sup> the last paper that reported the event occurred after multiple therapeutic procedures. The incidence of TACE-related deaths tended to decrease year by year. In the current study, hepatic failure (40.1%) ranked first, followed by cancer death (18.2%) and rupture of HCC (15.9%). Forty-five percent of 38 patients whose degree of liver damage was mentioned were in liver damage grade C. One study reported hepatic failure as the main adverse event of this therapy.<sup>26</sup> The postembolization rupture of HCC was uncommon but is life-threatening.<sup>51</sup>

As a limitation of this study, we did not analyze the relationship between the survival and dosage of the anticancer agent and lipiodol administered for TACE because of lack of items in the questionnaire sheet. It was presumed that doxorubicin<sup>20,21,23,28,32,52</sup> and epirubicin<sup>52</sup> analog of doxorubicin followed by cisplatin<sup>20,28</sup> were mainly used. The dose of lipiodol used was speculated to be approximately 5 mL/body and at maximum no more than 10 mL/body to prevent deterioration of hepatic function, which were mainly based on the recommended safety dose of lipiodol of 0.1 mL/kg<sup>53</sup> and 0.4 mL/kg<sup>54</sup> that were studied in normal animal liver combined with and without gelatin sponge particles, respectively. We failed to clarify which agent was more effective and how much of lipiodol was optimal.

As a main cause of a short mean follow-up period of 1.77 years, the following 2 factors can be given. One was early death of patients and the other was early increase of censored cases. During the first 2 years of follow-up,

66.1% (2383/3605) patients died, and 65.2% (3197/4905) patients were censored. The reason why the censored cases were high could be speculated as follows: the patients would frequently change hospitals to seek more specialized therapy, which was optimized to their stage of HCC and liver function. If these censored patients were low in frequency, the survival rates of TACE would have been probably lower than those of the current study because the majority of HCC patients has a poor prognosis. The high incidence of censored cases might reflect the limitation of the large scale of the current registration system. A new registration system has just started in which the drawbacks are improved, and the more precise and flexible information would be transmitted to the world.

In conclusion, the current study revealed that TACE demonstrated a high 5-year survival rate of 26% and a low TACE-related mortality rate of .5% in 8510 patients with unresectable HCC. The measurements of hepatic function (degree of liver damage proposed by the Liver Cancer Study Group of Japan), tumor characteristics (size, number, and presence of portal vein tumor thrombus evaluated by TNM staging), and AFP value were independent predictors of survival of patients. We hope these results will aid in predicting prognosis of a candidate patient who will be treated with the current procedure of TACE, in comparing the effectiveness with newly designed transarterial therapy, and in constructing a more practical new treatment algorithm.

## References

1. Parkin DM, Bray F, Ferlay J, Pisani P. Estimating the world cancer burden: Globocan 2000. *Int J Cancer* 2001;94:153-156.
2. Imamura H, Matsuyama Y, Tanaka E, Ohkubo T, Hasegawa K, Miyagawa S, Sugawara Y, Minagawa M, Takayama T, Kawasaki S, Makuuchi M. Risk factors contributing to early and late phase intrahepatic recurrence of hepatocellular carcinoma after hepatectomy. *J Hepatol* 2003;38:200-207.
3. Chen MF, Hwang TL, Jeng LB, Wang CS, Jan YY, Chen SC. Postoperative recurrence of hepatocellular carcinoma. Two hundred five consecutive patients who underwent hepatic resection in 15 years. *Arch Surg* 1994;129:738-742.
4. Koda M, Murawaki Y, Mitsuda A, Ohyama K, Horie Y, Suou T, Kawasaki H, Ikawa S. Predictive factors for intrahepatic recurrence after percutaneous ethanol injection therapy for small hepatocellular carcinoma. *Cancer* 2000;88:529-537.
5. Arai S, Yamaoka Y, Futagawa S, Inoue K, Kobayashi K, Kojiro M, Makuuchi M, Nakamura Y, Okita K, Yamada R. Results of surgical and nonsurgical treatment for small-sized hepatocellular carcinomas: a retrospective and nationwide survey in Japan. The Liver Cancer Study Group of Japan. *Hepatology* 2000;32:1224-1229.
6. Lin DY, Liaw YF, Lee TY, Lai CM. Hepatic arterial embolization in patients with unresectable hepatocellular carcinoma—a randomized controlled trial. *Gastroenterology* 1988;94:453-456.
7. Pelletier G, Roche A, Ink O, Anciaux ML, Derhy S, Rougier P, Lenoir C, Attali P, Etienne JP. A randomized trial of hepatic arterial chemoembolization in patients with unresectable hepatocellular carcinoma. *J Hepatol* 1990;11:181-184.

8. GETCH. A comparison of lipiodol chemoembolization and conservative treatment for unresectable hepatocellular carcinoma. *N Engl J Med* 1995;332:1256–1261.
9. Bruix J, Llovet JM, Castells A, Montana X, Bru C, Ayuso MC, Vilana R, Rodes J. Transarterial embolization versus symptomatic treatment in patients with advanced hepatocellular carcinoma: results of a randomized, controlled trial in a single institution. *Hepatology* 1998;27:1578–1583.
10. Pelletier G, Ducreux M, Gay F, Luboinski M, Hagege H, Dao T, Van Steenberg W, Buffet C, Rougier P, Adler M, Pignon JP, Roche A. Treatment of unresectable hepatocellular carcinoma with lipiodol chemoembolization: a multicenter randomized trial. *Group CHC. J Hepatol* 1998;29:129–134.
11. Llovet JM, Real MI, Montana X, Planas R, Coll S, Aponte J, Ayuso C, Sala M, Muchart J, Sola R, Rodes J, Bruix J. Arterial embolisation or chemoembolisation versus symptomatic treatment in patients with unresectable hepatocellular carcinoma: a randomised controlled trial. *Lancet* 2002;359:1734–1739.
12. Lo CM, Ngan H, Tso WK, Liu CL, Lam CM, Poon RT, Fan ST, Wong J. Randomized controlled trial of transarterial lipiodol chemoembolization for unresectable hepatocellular carcinoma. *Hepatology* 2002;35:1164–1171.
13. Doyon DMA, Jourde AM, Regensberg C, Frileux C. L'embolisation arterielle hepaticque dans les tumeurs malignes du foie. *Ann Radiol* 1974;17:593–603.
14. Yamada R, Nakatsuka H, Nakamura K, Sato M, Tamaoka K, Takemoto K, Minakuchi K, Yamaguchi S, Tamaki M, Monna T, Yamamoto S. Transcatheter arterial embolization therapy in unresectable hepatomas—experience in 15 cases. *Acta Hepatologica Japonica* 1979;20:595–603.
15. Konno T, Maeda H, Iwai K, Tashiro S, Maki S, Morinaga T, Mochinaga M, Hiraoka T, Yokoyama I. Effect of arterial administration of high-molecular-weight anticancer agent SMANCS with lipid lymphographic agent on hepatoma: a preliminary report. *Eur J Cancer Clin Oncol* 1983;19:1053–1065.
16. Ohishi H, Uchida H, Yoshimura H, Ohue S, Ueda J, Katsuragi M, Matsuo N, Hosogi Y. Hepatocellular carcinoma detected by iodized oil. Use of anticancer agents. *Radiology* 1985;154:25–29.
17. Yumoto Y, Jinno K, Tokuyama K, Araki Y, Ishimitsu T, Maeda H, Konno T, Iwamoto S, Ohnishi K, Okuda K. Hepatocellular carcinoma detected by iodized oil. *Radiology* 1985;154:19–24.
18. Takayasu K, Shima Y, Muramatsu Y, Moriyama N, Yamada T, Makuuchi M, Hasegawa H, Hirohashi S. Hepatocellular carcinoma: treatment with intra-arterial iodized oil with and without chemotherapeutic agents. *Radiology* 1987;163:345–351.
19. Sasaki Y, Imaoka S, Kasugai H, Fujita M, Kawamoto S, Ishiguro S, Kojima J, Ishikawa O, Ohigashi H, Furukawa H, et al. A new approach to chemoembolization therapy for hepatoma using ethiodized oil, cisplatin, and gelatin sponge. *Cancer* 1987;60:1194–1203.
20. Kasugai H, Kojima J, Tatsuta M, Okuda S, Sasaki Y, Imaoka S, Fujita M, Ishiguro S. Treatment of hepatocellular carcinoma by transcatheter arterial embolization combined with intraarterial infusion of a mixture of cisplatin and ethiodized oil. *Gastroenterology* 1989;97:965–971.
21. Nakamura H, Hashimoto T, Oi H, Sawada S. Transcatheter oily chemoembolization of hepatocellular carcinoma. *Radiology* 1989;170:783–786.
22. Uchida H, Ohishi H, Matsuo N, Nishimine K, Ohue S, Nishimura Y, Maeda M, Yoshioka T. Transcatheter hepatic segmental arterial embolization using lipiodol mixed with an anticancer drug and Gelfoam particles for hepatocellular carcinoma. *Cardiovasc Intervent Radiol* 1990;13:140–145.
23. Matsui O, Kadoya M, Yoshikawa J, Gabata T, Arai K, Demachi H, Miyayama S, Takashima T, Unoura M, Kogayashi K. Small hepatocellular carcinoma: treatment with subsegmental transcatheter arterial embolization. *Radiology* 1993;188:79–83.
24. Kawai S, Tani M, Okamura J, Ogawa M, Ohashi Y, Monden M, Hayashi S, Inoue J, Kawarada Y, Kusano M, Kubo Y, Kuroda C, Sakata Y, Shimamura Y, Jinno K, Takahashi A, Takayasu K, Tamura K, Nagasue N, Nakanishi Y, Makino M, Masuzawa M, Yumoto Y, Mori T, Oda T, The Cooperative Study Group for Liver Cancer Treatment of Japan. Prospective and randomized clinical trial for the treatment of hepatocellular carcinoma—a comparison between L-TAE with farmorubicin and L-TAE with adriamycin: preliminary results (second cooperative study). *Cancer Chemother Pharmacol* 1994;33(Suppl):S97–S102.
25. Chang JM, Tzeng WS, Pan HB, Yang CF, Lai KH. Transcatheter arterial embolization with or without cisplatin treatment of hepatocellular carcinoma. A randomized controlled study. *Cancer* 1994;74:2449–2453.
26. Bronowicki JP, Vetter D, Dumas F, Boudjema K, Bader R, Weiss AM, Wenger JJ, Boissel P, Bigard MA, Doffoel M. Transcatheter oily chemoembolization for hepatocellular carcinoma. A 4-year study of 127 French patients. *Cancer* 1994;74:16–24.
27. Stefanini GF, Amorati P, Biselli M, Mucci F, Celi A, Arienti V, Roversi R, Rossi C, Re G, Gasbarrini G. Efficacy of transarterial targeted treatments on survival of patients with hepatocellular carcinoma. An Italian experience. *Cancer* 1995;75:2427–2434.
28. Hatanaka Y, Yamashita Y, Takahashi M, Koga Y, Saito R, Nakashima K, Urata J, Miyao M. Unresectable hepatocellular carcinoma: analysis of prognostic factors in transcatheter management. *Radiology* 1995;195:747–752.
29. Chung JW, Park JH, Han JK, Choi BI, Han MC, Lee HS, Kim CY. Hepatic tumors: predisposing factors for complications of transcatheter oily chemoembolization. *Radiology* 1996;198:33–40.
30. Solomon B, Soulen MC, Baum RA, Haskal ZJ, Shlansky-Goldberg RD, Cope C. Chemoembolization of hepatocellular carcinoma with cisplatin, doxorubicin, mitomycin-C, ethiodol, and polyvinyl alcohol: prospective evaluation of response and survival in a US population. *J Vasc Interv Radiol* 1999;10:793–798.
31. Savastano S, Miotto D, Casarrubea G, Teso S, Chiesura-Corona M, Feltrin GP. Transcatheter arterial chemoembolization for hepatocellular carcinoma in patients with Child's grade A or B cirrhosis: a multivariate analysis of prognostic factors. *J Clin Gastroenterol* 1999;28:334–340.
32. Takayasu K, Muramatsu Y, Maeda T, Iwata R, Furukawa H, Moriyama N, Okusaka T, Okada S, Ueno H. Targeted transarterial oily chemoembolization for small foci of hepatocellular carcinoma using a unified helical CT and angiography system: analysis of factors affecting local recurrence and survival rates. *AJR Am J Roentgenol* 2001;176:681–688.
33. Ikai I, Arii S, Ichida T, Okita K, Omata M, Kojiro M, Takayasu K, Nakanuma Y, Makuuchi M, Matsuyama Y, Yamaoka Y. Report of the 16th follow-up survey of primary liver cancer. *Hepatol Res* 2005;32:163–172.
34. Araki T, Itai Y, Furui S, Tasaka A. Dynamic CT densitometry of hepatic tumors. *AJR Am J Roentgenol* 1980;135:1037–1043.
35. The Liver Cancer Study Group of Japan. The general rules for the clinical and pathological study of primary liver cancer. Version 4. Tokyo: Kanehara & Co., Ltd., 2000.
36. The Liver Cancer Study Group of Japan. The general rules for the clinical and pathological study of primary liver cancer. Version 2. Tokyo: Kanehara & Co., Ltd., 2003.
37. Makuuchi M, Belghiti J, Belli G, Fan ST, Lau JW, Ringe B, Strasberg SM, Vauthey JN, Yamaoka Y, Yamasaki S. IHPBA concordant classification of primary liver cancer: working group report. *J Hepatobiliary Pancreat Surg* 2003;10:26–30.
38. Yamada R, Kishi K, Sonomura T, Tsuda M, Nomura S, Satoh M. Transcatheter arterial embolization in unresectable hepatocellular carcinoma. *Cardiovasc Intervent Radiol* 1990;13:135–139.
39. Ikeda K, Kumada H, Saitoh S, Arase Y, Chayama K. Effect of repeated transcatheter arterial embolization on the survival time

- in patients with hepatocellular carcinoma. An analysis by the Cox proportional hazard model. *Cancer* 1991;68:2150–2154.
40. The Liver Cancer Study Group of Japan. Report of the 12th follow-up survey of primary liver cancer. Kyoto: Shinkou-insatsu Press, 1996:100.
  41. Trevisani F, D'Intino PE, Caraceni P, Pizzo M, Stefanini GF, Mazziotti A, Grazi GL, Gozzetti G, Gasbarrini G, Bernardi M. Etiologic factors and clinical presentation of hepatocellular carcinoma. Differences between cirrhotic and noncirrhotic Italian patients. *Cancer* 1995;75:2220–2232.
  42. The Liver Cancer Study Group of Japan. Predictive factors for long term prognosis after partial hepatectomy for patients with hepatocellular carcinoma in Japan. *Cancer* 1994;74:2772–2780.
  43. Vauthey JN, Lauwers GY, Esnaola NF, Do KA, Beighiti J, Mirza N, Curley SA, Ellis LM, Regimbeau JM, Rashid A, Cleary KR, Nagorney DM. Simplified staging for hepatocellular carcinoma. *J Clin Oncol* 2002;20:1527–1536.
  44. Okuda K, Ohtsuki T, Obata H, Tomimatsu M, Okazaki N, Hasegawa H, Nakajima Y, Ohnishi K. Natural history of hepatocellular carcinoma and prognosis in relation to treatment. Study of 850 patients. *Cancer* 1985;56:918–928.
  45. Llovet JM, Bru C, Bruix J. Prognosis of hepatocellular carcinoma: the BCLC staging classification. *Semin Liver Dis* 1999;19:329–338.
  46. 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. *Hepatology* 2000;31:840–845.
  47. Kudo M, Chung H, Haji S, Osaki Y, Oka H, Seki T, Kasugai H, Sasaki Y, Matsunaga T. Validation of a new prognostic staging system for hepatocellular carcinoma: the JIS score compared with the CLIP score. *Hepatology* 2004;40:1396–1405.
  48. Nanashima A, Sumida Y, Morino S, Yamaguchi H, Tanaka K, Shibasaki S, Ide N, Sawai T, Yasutake T, Nakagoe T, Nagayasu T. The Japanese integrated staging score using liver damage grade for hepatocellular carcinoma in patients after hepatectomy. *Eur J Surg Oncol* 2004;30:765–770.
  49. Ernst O, Sergeant G, Mizrahi D, Delemazure O, Paris JC, L'Hermine C. Treatment of hepatocellular carcinoma by transcatheter arterial chemoembolization: comparison of planned periodic chemoembolization and chemoembolization based on tumor response. *AJR Am J Roentgenol* 1999;172:59–64.
  50. Vetter D, Wenger JJ, Bergier JM, Doffoel M, Bockel R. Transcatheter oily chemoembolization in the management of advanced hepatocellular carcinoma in cirrhosis: results of a Western comparative study in 60 patients. *Hepatology* 1991;13:427–433.
  51. Nakao N, Kotake M, Miura K, Ohnishi M, Miura T. Rupture of hepatocellular carcinoma following transcatheter arterial embolization. *Radiat Med* 1988;6:147–149.
  52. Kawai S, Tani M, Okamura J, Ogawa M, Ohashi Y, Monden M, Hayashi S, Inoue J, Kawarada Y, Kusano M, Kubo Y, Kuroda C, Sakata Y, Shimamura Y, Jinno K, Takahashi A, Takayasu K, Tamura K, Nagasue N, Nakanishi Y, Makino M, Masuzawa M, Yumoto Y, Mori T, Oda T. Prospective and randomized trial of lipiodol-transcatheter arterial chemoembolization for treatment of hepatocellular carcinoma: a comparison of epirubicin and doxorubicin (second cooperative study). The Cooperative Study Group for Liver Cancer Treatment of Japan. *Semin Oncol* 1997;24:S6–S45.
  53. Sato M, Yamada R, Uchida B, Hedgepeth P, Rosch J. Effects of hepatic artery embolization with Lipiodol and gelatin sponge particles on normal swine liver. *Cardiovasc Intervent Radiol* 1993;16:348–354.
  54. Kan Z, Wallace S. Sinusoidal embolization: impact of iodized oil on hepatic microcirculation. *J Vasc Interv Radiol* 1994;5:881–886.

---

Received August 21, 2005. Accepted April 27, 2006.

Address requests for reprints to: Kenichi Takayasu, MD, Department of Diagnostic Radiology, National Cancer Center Hospital, 5-1-1, Tsukiji, Chuo-ku, Tokyo 104-0045, Japan. e-mail: ktakayas@ncc.go.jp; fax: (81) 3-3542-3815.

## A modified Japan Integrated Stage score for prognostic assessment in patients with hepatocellular carcinoma

IWAO IKAI<sup>1</sup>, KENICHI TAKAYASU<sup>2</sup>, MASAO OMATA<sup>3</sup>, KIWAMU OKITA<sup>4</sup>, YASUNI NAKANUMA<sup>5</sup>, YUTAKA MATSUYAMA<sup>6</sup>, MASATOSHI MAKUUCHI<sup>7</sup>, MASAMICHI KOJIRO<sup>8</sup>, TAKAFUMI ICHIDA<sup>9</sup>, SHIGEKI ARII<sup>10</sup>, and YOSHIO YAMAOKA<sup>1</sup> for the Liver Cancer Study Group of Japan

<sup>1</sup>Department of Gastroenterological Surgery, Kyoto University Graduate School of Medicine, Kyoto, Japan

<sup>2</sup>Department of Diagnostic Radiology, National Cancer Center Hospital, Tokyo, Japan

<sup>3</sup>Department of Gastroenterology, Graduate School of Medicine, University of Tokyo, Tokyo, Japan

<sup>4</sup>Department of Gastroenterology and Hepatology, Yamaguchi University School of Medicine, Ube, Japan

<sup>5</sup>Department of Human Pathology, Kanazawa University Graduate School of Medicine, Kanazawa, Japan

<sup>6</sup>Department of Biostatistics, School of Health Sciences and Nursing University of Tokyo, Tokyo, Japan

<sup>7</sup>Hepato-Biliary-Pancreatic Surgery Division, Department of Surgery, Graduate School of Medicine, University of Tokyo, Tokyo, Japan

<sup>8</sup>Department of Pathology, Kurume University School of Medicine, Kurume, Japan

<sup>9</sup>Department of Gastroenterology, Juntendo University School of Medicine, Tokyo, Japan

<sup>10</sup>Department of Hepato-Biliary-Pancreatic Surgery, Tokyo Medical and Dental University Graduate School of Medicine, Tokyo, Japan

**Background.** For patients with hepatocellular carcinoma (HCC), the selection of effective therapeutic options depends on a reliable prognostic assessment of both the tumor characteristics and liver impairment. This study sought to provide a modified Japan Integrated Stage (m-JIS) score to predict more accurately the survival of patients with HCC. **Methods.** We analyzed the records of 42269 patients diagnosed with HCC that were registered between 1992 and 1999 in a nationwide Japanese database. The m-JIS score was calculated from the tumor-node-metastasis stage and the grade of liver damage as defined by the Liver Cancer Study Group of Japan. The predictive accuracy for patient survival based on the m-JIS score was compared with that determined by the modified Cancer of the Liver Italian Program (m-CLIP) score using the cross-validation method. **Results.** Patients were divided randomly into two groups, a training sample for construction of prediction models ( $n = 21\,127$  patients with 8458 deaths) and a validation sample for assessment of those prediction models ( $n = 21\,142$  patients with 8434 deaths). Both the m-JIS score and the m-CLIP score showed good discriminatory ability in the training sample. In the validation sample, the residuals of prediction based on the m-JIS score were smaller than those of the m-CLIP score ( $P < 0.0001$ ). **Conclusions.** The m-JIS scoring system had better predictive accuracy than the m-CLIP score system for survival of Japanese patients with HCC.

**Key words:** TNM Stage, degree of liver damage, nationwide survey database, Liver Cancer Study Group of Japan, CLIP score

### Introduction

Hepatocellular carcinoma (HCC) is one of the most common malignancies worldwide. For HCC patients, several effective treatment modalities exist, including liver resection, local ablation therapy, transcatheter arterial chemoembolization, and liver transplantation. However, the efficacy of these modalities is limited by both tumor status and liver dysfunction, because most HCC patients present with chronic hepatitis or liver cirrhosis due to hepatitis B or C viral infection.<sup>1</sup> In Japan, more than 85% of HCC patients have hepatitis B or C viral infection.<sup>2</sup> Therefore, a reliable prognostic assessment based on tumor characteristics and the grade of liver impairment is required to predict outcome in patients.

During the last decade, better screening of high-risk populations with hepatitis virus B or C infection has led to increases in the early detection of HCC. Several advances in diagnostic techniques and therapeutic modalities have significantly improved the prognosis of patients with HCC.<sup>3–5</sup> In Japan, the 5-year survival rate of patients with HCC who underwent liver resection increased from less than 40% during the 1980s to greater than 50% during the 1990s.<sup>6</sup> We have reported that age, degree of liver damage,  $\alpha$ -fetoprotein (AFP) levels, maximum tumor diameter, number of tumors, presence of portal vein or hepatic vein invasion, extrahepatic metastasis, surgical curability, and surgical free margin are independent prognostic factors for HCC

Received: October 11, 2005 / Accepted: July 25, 2006

Reprint requests to: I. Ikai

Present address: The Liver Cancer Study Group of Japan, 403 Bear House, 40 Sanno-cho, Shogoin, Sakyo-ku, Kyoto 606-8392, Japan

patients who undergo liver resection.<sup>6</sup> These findings prompted the revision of the tumor-node-metastasis (TNM) staging system by the Liver Cancer Study Group of Japan (LCSGJ).<sup>7</sup> In this system, tumor status is determined by maximum tumor diameter, total number of tumors, and the presence of vascular or bile duct invasion. To evaluate liver function precisely, the LCSGJ also proposed a classification of the degree of liver damage to replace the Child-Pugh classification system.<sup>7</sup> Both TNM staging and the LCSGJ degree of liver damage classification are useful for predicting the prognosis of patients with HCC.

Staging systems are created to predict the outcome of patients and to evaluate the efficacy of therapeutic modalities. Such systems are important research tools, permitting comparison of different groups and trials.<sup>8</sup> Recently, to predict outcome in patients with HCC, new HCC staging systems combining tumor status with the degree of liver damage, namely, the Cancer of the Liver Italian Program (CLIP) score,<sup>9</sup> the Barcelona Clinic Liver Cancer (BCLC) staging classification,<sup>10</sup> the Chinese University Prognostic Index (GUPI),<sup>11</sup> and the Japan Integrated Stage (JIS) score,<sup>12,13</sup> have been proposed.

In this study, we propose a new staging system, defined by the LCSGJ TNM stage and the LCSGJ degree of liver damage classification. This staging system is a modified JIS score that incorporates the degree of liver damage classification, instead of the Child-Pugh classification, to evaluate liver function more precisely. We calculated the modified JIS (m-JIS) score of patients in a nationwide HCC survey database and compared the predictive accuracy of the m-JIS score for the survival of patients with that of a modified CLIP (m-CLIP) score.

## Patients and methods

In a nationwide follow-up survey of primary liver cancer conducted by the LCSGJ, patients from approximately 800 institutions in Japan with primary malignant liver tumors diagnosed by imaging studies, preoperative clinical data, or histopathological studies were registered every 2 years. Registered patients were followed up prospectively. In this study, we collected the records of HCC patients who were registered in the database between January 1992 and December 1999 and for whom complete information on six tumor characteristics in imaging studies (maximum tumor diameter, number of tumors, portal and hepatic vein invasion, and lymph node or distant organ metastasis) for determination of the TNM stage, degree of liver damage, and AFP levels at diagnosis was available.

TNM stage and the degree of liver damage classification are defined by the LCSGJ in the General Rules for the Clinical and Pathological Study of Primary Liver Cancer, second English edition (Tables 1 and 2).<sup>7</sup> In this study, the TNM stage was determined by tumor characteristics in imaging studies, not from resected specimens, even when hepatic resection was performed. Bile duct invasion in imaging studies was excluded from our determination of TNM stage owing to the lack of a corresponding item in the database. The m-JIS score was calculated using the TNM stage and the grade of liver damage (Table 3). The original CLIP score was determined from tumor morphology, AFP levels, portal vein invasion, and liver function. Unfortunately, the Child-Pugh classification and tumor morphology with "extent >50% of the liver" for calculating the original CLIP score were not included in our database. In the m-

**Table 1.** The TNM classification proposed by LCSGJ

Factors	Factor score	Determination
T factor		
1. Solitary tumor	T1	Meets all of the criteria
2. Maximum tumor diameter $\leq 2$ cm	T2	Meets two of the three criteria
3. No vascular invasion (to PV, HV, or BD)	T3	Meets one of the three criteria
	T4	Does not meet any of the criteria
N factor		
Lymph node metastasis	N0	Absent
	N1	Present
M factor		
Distant metastasis	M0	Absent
	M1	Present
Stage		
I		T1 N0 M0
II		T2 N0 M0
III		T3 N0 M0
IV-A		T4 N0 M0; any T N1 M0
IV-B		Any T or N M1

TNM, tumor-node-metastasis; LCSGJ, Liver Cancer Study Group of Japan; PV, portal vein; HV, hepatic vein; BD, bile duct

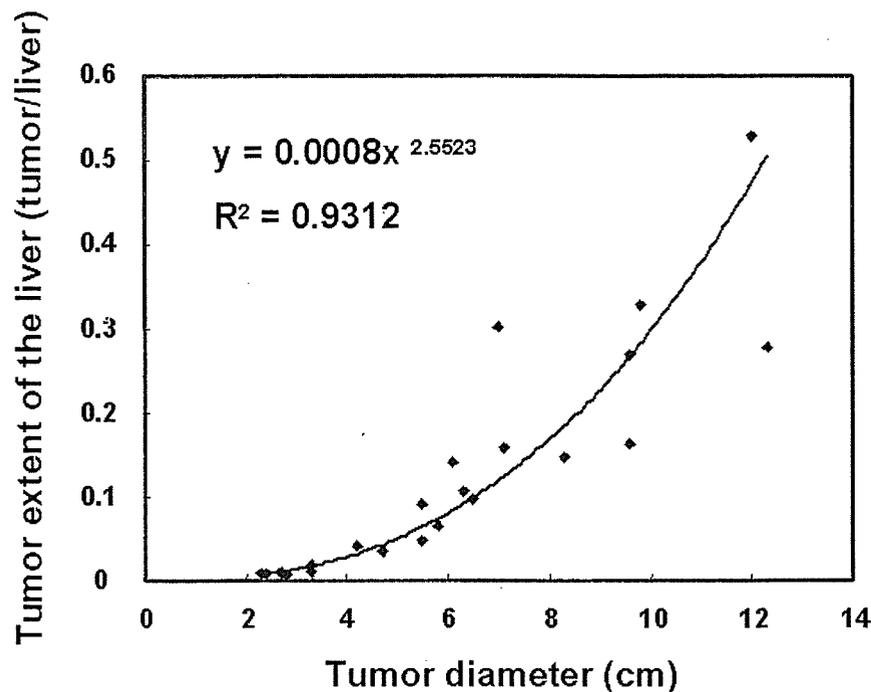


Fig. 1. Relationship between maximum tumor diameter and tumor extent in 23 patients with hepatocellular carcinoma

Table 2. Degree of liver damage classification by LCSGJ

Item	Liver damage grade		
	A	B	C
Ascites	None	Controllable	Uncontrollable
Serum bilirubin (mg/dl)	<2.0	2.0–3.0	>3.0
Serum albumin (g/dl)	>3.5	3.0–3.5	>3.0
ICGR <sub>15</sub> (%)	<15	15–40	>40
Prothrombin activity (%)	>80	50–80	<50

The highest grade with at least two items meeting the criteria for that grade is adopted as the degree of liver damage

ICGR<sub>15</sub>, indocyanine green retention rate at 15 min

Table 3. The modified Japan Integrated Stage (m-JIS) scoring system

Liver damage grade	TNM stage	Value assigned
A	I	0
B	II	1
C	III	2
	IV	3

m-JIS score = (liver damage grade) + (TNM stage)

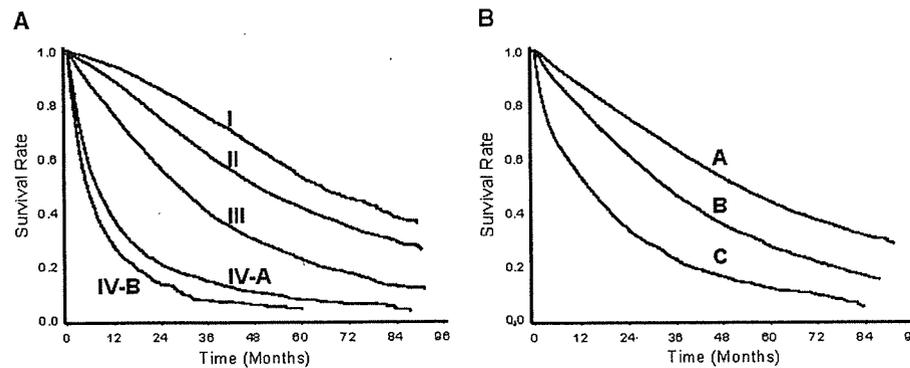
CLIP score, the grade of liver damage, indicated in Table 2, was applied for liver function, instead of the Child-Pugh classification. In a preliminary study, we examined the relationship between maximum tumor diameter and tumor extent in 23 patients with HCC who underwent computed tomography at Kyoto University Hospital. “Extent >50% of the liver” corresponded to a

tumor diameter of approximately 12 cm (Fig. 1). In the nationwide database, maximum tumor diameter groups were 0–2, 2–3, 3–5, 5–10, 10–15, 15–20, 20–25, and more than 25 cm. In this study, tumor morphology with “extent >50% of the liver” was changed to maximum tumor diameter >10 cm (Table 4). In the original CLIP score, 400 ng/dl was defined as the cutoff level for AFP;<sup>9</sup> we used 400 ng/ml as the cutoff level in the m-CLIP score, as only 6.2% of patients had AFP levels above 10000 ng/ml.

Overall cumulative survival rates were obtained with the Kaplan-Meier method. To calculate survival times, the starting point was the date of admission and the end point was the date of either death or the last follow-up information. Follow-up ended on 31 December 1999. The mean follow-up period was 23.2 months (range, 1–96 months). Differences in survival between groups were compared by a log rank test. To evaluate the m-JIS

**Table 4.** The modified Cancer of the Liver Italian Program (m-CLIP) scoring system

Degree of liver damage	Tumor morphology	AFP (ng/ml)	Portal thrombosis	Value assigned
A	Uninodular and maximum tumor diameter $\leq 10$ cm	$\leq 400$	No	0
B	Multinodular and maximum tumor diameter $\leq 10$ cm	$> 400$	Yes	1
C	Maximum tumor diameter $> 10$ cm			2

AFP,  $\alpha$ -fetoprotein**Fig. 2.** Survival curve of patients with hepatocellular carcinoma according to the tumor-node-metastasis (TNM) stage (A) and degree of liver damage classification (B) by the Liver Cancer Study Group of Japan. We observed statistically significant differences among groups (stages or liver damage grades) for both TNM stage and the degree of liver damage classification ( $P < 0.0001$ )

score as a predictor, independent variables [m-JIS score, age, sex, hepatitis B virus surface antigen (HBsAg), anti-hepatitis C virus antibody (HCVAb), tumor size, number of tumors, portal and hepatic vein invasion, extrahepatic metastasis, and AFP level] were used in a multivariate analysis by Cox's proportional hazard model.

The predictive accuracies of the m-JIS score and the m-CLIP score for the survival of patients were compared using the cross-validation method. For cross validation, patients were divided randomly into two groups: a training sample and a validation sample. In the training sample, two predictive models that included either the m-JIS or the m-CLIP score as a covariate were constructed using Cox's proportional hazards model. In the validation sample, each estimated model from the test sample was applied to predict the survival of each patient. The predicted survival curves, plotted for each model, were compared to the observed survival curves of the validation sample, plotted by the Kaplan-Meier method. The predictive accuracy was compared in terms of the residual, the difference between the observed survival time and the predicted survival time in the validation sample. Small absolute values of the residuals indicate accurate prediction of the survival of patients. The estimated absolute values of the residuals were compared between the m-JIS and m-CLIP scores by the analysis of variance (ANOVA) method. For the ANOVA, the generalized estimating equation (GEE)

approach was used to account for the correlation between the two residuals in each patient.<sup>14</sup>

$P$  values less than 0.05 were considered statistically significant. Statistical analysis was carried out with SAS ver. 8.02.

## Results

Records of a total of 42269 registered patients were collected from the database. The patients' profiles are shown in Table 5. The average patient age was  $64.1 \pm 9.3$  (mean  $\pm$  SD), and the ratio of men to women was 2.9:1. Among total patients, 15.4% tested positive for HBsAg and 74.6% were positive for HCVAb. According to the imaging studies, 27.0% of the patients had tumor(s) with a maximum diameter of 2 cm or less, and 76.9% of 5 cm or less. Only 7.6% of the patients had a tumor with maximum diameter greater than 10 cm. More than half of the patients had a solitary tumor. Portal and hepatic vein invasions were observed in 12.4% and 4.7% of the patients, respectively. In 3.7% of all patients, extrahepatic metastasis, including lymph node and distant organ metastases, was observed. As the initial treatment for HCC, 13119 patients underwent liver resection, 11302 received local ablation therapy, 13868 underwent transcatheter arterial chemoembolization, and 3980 patients received other treatments or best supportive therapy. Up through 1999, no patients who underwent

**Table 5.** Patient profiles and tumor characteristics

	<i>n</i>	(%)
Age (mean ± SD)		
64.1 ± 9.3		
Sex		
Male	31 537	74.6
Female	10 732	25.4
HBsAg		
Negative	34 627	81.9
Positive	6 523	15.4
NR	1 119	2.7
HCVAb		
Negative	9 956	23.6
Positive	31 548	74.6
NR	765	1.8
Tumor size		
Diameter ≤2 cm	11 426	27.0
2 cm < diameter ≤3 cm	10 683	25.3
3 cm < diameter ≤5 cm	10 412	24.6
5 cm < diameter ≤10 cm	6 552	15.5
10 cm < diameter	3 196	7.6
Number of tumors		
Solitary	24 044	56.9
Multiple	18 225	43.1
Portal vein invasion		
Absent	37 013	87.6
Present	5 256	12.4
Hepatic vein invasion		
Absent	40 299	95.3
Present	1 970	4.7
Extrahepatic metastasis		
Absent	40 693	96.3
Present	1 576	3.7
AFP (ng/ml)		
≤400	32 421	76.7
>400	9 848	23.3
TNM Stage		
I	7 439	17.6
II	17 998	42.6
III	12 571	29.7
IV-A	3 023	7.2
IV-B	1 238	2.9
Degree of liver damage		
A	22 732	53.8
B	15 152	35.9
C	4 385	10.4
m-JIS score		
0	4 698	11.1
1	12 475	29.5
2	13 265	31.4
3	7 591	18.0
4	3 133	7.4
5	1 107	2.6
m-CLIP score		
0	10 775	25.5
1	13 838	32.7
2	9 733	23.0
3	4 690	11.1
4	1 997	4.7
5	897	2.1
6	339	0.8

HBsAg, hepatitis B virus surface antigen; HCVAb, anti-hepatitis C virus antibody; NR, not recorded

liver transplantation for HCC were registered in our database.

The cumulative survival curves of patients according to the TNM stage indicated that the 5-year survival rates were 53.0%, 41.3%, 22.1%, 7.0%, and 3.2% in patients with stage I ( $n = 7439$ ), stage II ( $n = 17998$ ), stage III ( $n = 12571$ ), stage IV-A ( $n = 3023$ ), and stage IV-B ( $n = 1238$ ) disease, respectively (Fig. 2A). The 5-year survival rates were 43.4% for liver damage grade A ( $n = 22732$ ), 26.7% for grade B ( $n = 15152$ ), and 11.2% for grade C ( $n = 4385$ ) (Fig. 2B). There were statistically significant differences in survival among groups, whether calculated for TNM stages or the degree of liver damage ( $P < 0.0001$ ).

Multivariate analysis with Cox's proportional hazard model indicated that the m-JIS score, patient age, HBsAg, tumor size, number of tumors, portal vein invasion, hepatic vein invasion, extrahepatic metastasis, and AFP level were independent prognostic predictors for HCC patients (Table 6). Among these variables, the m-JIS score had the highest hazard ratio. Therefore, the predictive accuracy for patient survival based on the m-JIS score was compared with that determined by the m-CLIP score by using the cross-validation method.

The 42269 patients were divided randomly into two groups: a training sample ( $n = 21127$  patients with 8458 deaths) and a validation sample ( $n = 21142$  patients with 8434 deaths). Tables 7 and 8 show the prediction models constructed by Cox's proportional hazards method from the training sample based on the m-JIS and the m-CLIP score, respectively. In the training sample, both of these models had a good discriminating ability. Figure 3 shows the results of the application of each predictive model to the validation sample. The residual sum of squares was 17.49 for models based on the m-JIS score (Fig. 3A) and 21.81 for those based on the m-CLIP score (Fig. 3B) ( $P < 0.0001$ ). Table 9 details the results of the GEE analysis adjustment of the score values. The negative parameter estimate for the scoring method (m-JIS vs m-CLIP) indicates that the residuals based on the m-JIS score were smaller than those based on the m-CLIP score ( $P < 0.0001$ ).

## Discussion

To distinguish patients with a curable disease from those with more advanced disease, it is important to be able to assess the prognosis of HCC patients using imaging studies and clinical data, without a histopathological review of a resected specimen. The Okuda staging classification proposed in the 1980s is a simple staging system combining tumor characteristics with liver function.<sup>15</sup> In this staging system, however, tumor character is evaluated only by whether the tumor has extended

**Table 6.** Prognostic variables by multivariate analysis using Cox's proportional hazard model

Variable		Hazard ratio	95% C.I.	P value
m-JIS	m-JIS 1 vs m-JIS 0	1.44	1.33, 1.56	0.0001
	m-JIS 2 vs m-JIS 0	2.17	2.00, 2.35	0.0001
	m-JIS 3 vs m-JIS 0	3.45	3.16, 3.77	0.0001
	m-JIS 4 vs m-JIS 0	4.99	4.52, 5.51	0.0001
	m-JIS 5 vs m-JIS 0	8	7.09, 9.03	0.0001
Age (years)	≥60 vs <60	1.12	1.08, 1.16	0.0001
Sex	Male vs female	1.01	0.98, 1.05	0.54
HBsAg	Positive vs negative	1.16	1.10, 1.22	0.0001
HCVAb	Positive vs negative	1.04	0.99, 1.09	0.09
Tumor size (cm)	>2.0 vs ≤2.0	1.1	1.08, 1.11	0.0001
Number of tumors	Multiple vs solitary	1.06	1.02, 1.11	0.002
Portal vein invasion	Present vs absent	1.48	1.41, 1.56	0.0001
Hepatic vein invasion	Present vs absent	1.31	1.23, 1.40	0.0001
Extrahepatic metastasis	Present vs absent	1.81	1.69, 1.94	0.0001
AFP (ng/ml)	>400 vs ≤400	1.61	1.55, 1.66	0.0001

C.I., confidence interval

**Table 7.** The predictive model based on the m-JIS score

Parameter	Estimate	Standard error	P value	Hazard ratio	95% C.I.
Score 1 (m-JIS 1 vs m-JIS 0)	0.544	0.055	<0.0001	1.72	1.55, 1.92
Score 2 (m-JIS 2 vs m-JIS 0)	1.06	0.053	<0.0001	2.89	2.60, 3.20
Score 3 (m-JIS 3 vs m-JIS 0)	1.712	0.054	<0.0001	5.54	4.89, 6.16
Score 4 (m-JIS 4 vs m-JIS 0)	2.328	0.058	<0.0001	10.3	9.15, 11.5
Score 5 (m-JIS 5 vs m-JIS 0)	3.375	0.068	<0.0001	29.2	25.6, 33.4

**Table 8.** The predictive model based on the m-CLIP score

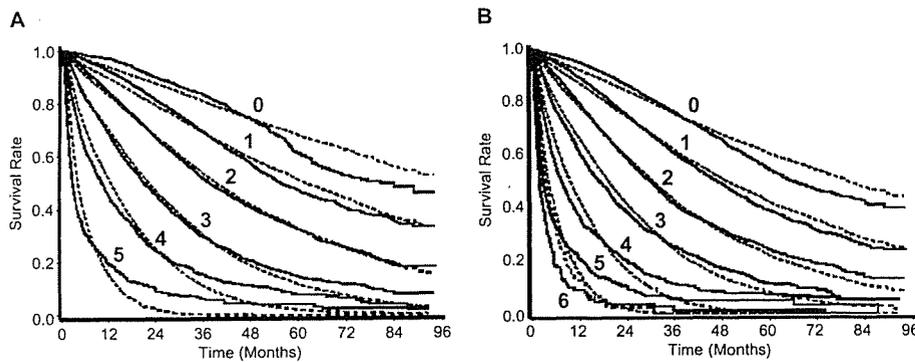
Parameter	Estimate	Standard error	P value	Hazard ratio	95% C.I.
Score 1 (m-CLIP 1 vs m-CLIP 0)	0.538	0.036	<0.0001	1.71	1.60, 1.84
Score 2 (m-CLIP 2 vs m-CLIP 0)	1.103	0.036	<0.0001	3.01	2.81, 3.24
Score 3 (m-CLIP 3 vs m-CLIP 0)	1.623	0.04	<0.0001	5.07	4.69, 5.48
Score 4 (m-CLIP 4 vs m-CLIP 0)	2.344	0.048	<0.0001	10.4	9.49, 11.4
Score 5 (m-CLIP 5 vs m-CLIP 0)	3.156	0.06	<0.0001	23.5	20.9, 26.4
Score 6 (m-CLIP 6 vs m-CLIP 0)	3.336	0.089	<0.0001	28.1	25.6, 33.4

**Table 9.** Generalized estimating equation analysis of the absolute value of the residuals

Parameter	Estimate	Standard error	P value
Intercept	0.0145	0.0002	<0.0001
Scoring method (m-JIS vs m-CLIP)	-0.0047	0.0001	<0.0001
Score value	0.0057	0.0001	<0.0001

throughout more than 50% of the liver. This criterion is used to determine not only the CLIP score but also the BCLC stage.<sup>9,10</sup> The assessment of tumor extension relative to total liver volume is poorly reproducible. In addition, tumor extension of more than 50% is indicative

of highly advanced disease. Because the recent promotion of screening programs of high-risk populations with hepatitis B or C viral infection has led to the increasingly early detection of HCC, the majority of tumors detected are small. In this study, 77% of patients had



**Fig. 3A,B.** Comparison of the observed survival curves (solid lines) and the predicted survival curves (dashed lines) in the validation sample ( $n = 21\,142$  patients). **A** Predictions based on the modified Japan Integrated Stage score; **B** those based on the modified Cancer of the Liver Italian Program score

tumors 5cm or less in diameter. Therefore, a precise evaluation of tumor character should be included in a staging system.

Recently, the TNM staging systems proposed by both the LCSGJ and the International Union Against Cancer/American Joint Committee on Cancer (UICC/AJCC) have been revised.<sup>7,16</sup> The new classification system designed by the LCSGJ is quite simple and convenient for clinical practitioners. Tumor status is determined by tumor size and number and by the presence or absence of vascular or bile duct invasion. The tumor size cutoff is 2cm in the LCSGJ TNM classification and 5cm in the UICC/AJCC classification. To facilitate more precise estimation of patient prognosis, the working group of the International Scientific Committee of the International Hepato-Pancreato-Biliary Association recommended the TNM staging system proposed by the LCSGJ.<sup>17</sup> In this study, the LCSGJ TNM stage had good ability to predict the survival of patients with HCC.

Kudo et al.<sup>12</sup> originally proposed the JIS score, which is defined by the LCSGJ TNM stage and the Child-Pugh classification. Using 4525 patients with HCC at five institutions, they validated the JIS score as a good prognostic staging system with good internal reproducibility.<sup>13</sup> Our study evaluated the m-JIS score, which extends the JIS score, using a large HCC population. The difference between these two systems is the scoring system for liver function. The Child-Pugh classification is used extensively internationally to evaluate liver function in patients with HCC. This classification was devised to predict the tolerance of patients with portal hypertension to portosystemic shunting. Encephalopathy is one of the variables in the Child-Pugh classification. In a recent report of the LCSGJ, mild to severe encephalopathy was observed in only 5.0% of all patients with HCC.<sup>2</sup> The indocyanine green (ICG) clearance test, one of the best laboratory tests to evaluate liver function precisely,<sup>18</sup> is widely used in Eastern countries. In the LCSGJ report, ICG retention rates at 15 min (ICGR<sub>15</sub>) were less than 15% in 33% of patients, 15%–40% in 54%, and more than 40% in 13%.<sup>2</sup> There-

fore, the LCSGJ proposed the degree of liver damage classification system, defined by ascites, serum bilirubin level, serum albumin level, ICGR<sub>15</sub>, and prothrombin activity, but excluding encephalopathy, to evaluate liver function precisely. To assess the prognosis of patients with HCC more effectively, we have modified the JIS score to incorporate the degree of liver damage classification instead of the Child-Pugh classification. The m-JIS score is calculated from TNM staging, determined using imaging studies, and the degree of liver damage classification, determined by clinical data. The m-JIS score indicated the most important predictor with the highest hazard ratio among several variables. Patient populations with m-JIS scores of 0–5 were well discriminated using this system.

The BCLC staging classification considers several critical prognostic factors, including performance status, tumor stage, Okuda stage, Child-Pugh classification, portal hypertension, and bilirubin levels.<sup>10</sup> The CUPJ system is another common prognostic staging system that utilizes similar prognostic indicators.<sup>11</sup> Unfortunately, we could not compare the BCLC and CUPJ staging systems with m-JIS staging, because portal hypertension, performance status, and alkaline phosphatase levels were not included in the original survey items within the database. The CLIP score proposed in 1998 is determined by Child-Pugh classification, tumor morphology, AFP levels, and portal vein thrombosis.<sup>9</sup> This score has been validated not only in multiple series using the same group<sup>19</sup> but also in Canadian and Japanese patients with HCC.<sup>20,21</sup>

In an evaluation of several HCC staging systems, Wildi et al.<sup>8</sup> reported that the CLIP score is the most well documented and analyzed staging system for prognostic purposes. In this study, we examined the m-CLIP score, based on a maximum tumor diameter greater than 10cm and the grade of liver damage by LCSGJ, because the original CLIP score could not be calculated with the information from the database, which did not include >50% tumor extension or the Child-Pugh classification. Also, we used 400ng/ml as the cutoff level for AFP in the m-CLIP score. The m-CLIP score may

be much better than the original CLIP score in its stratification ability and the survival prediction of each score, because the grade of liver damage by LCSGJ evaluates liver function more precisely than the Child-Pugh classification, and 400 ng/ml for AFP is a more useful cutoff level than 400 ng/dl in our patient population. The m-CLIP score also had a good discriminating ability for patients with HCC in the training sample.

In our comparison of the m-JIS score and the m-CLIP score using cross-validation and GEE analyses, the m-JIS score had a greater predicative accuracy than the m-CLIP score for the survival of patients with HCC. In both the m-JIS and m-CLIP analyses, only patients on whom ICG tests had been performed were included. In this study, 40.6% of total patients had an m-JIS score of 0 or 1 and 10.0% had an m-JIS score of 4 or 5, while 60.9% had an m-CLIP score of 0 or 1 and 7.6% had an m-CLIP score of 4, 5, or 6. This patient distribution is similar to that in the study of Kudo et al.<sup>13</sup> The difference in patient distribution might be an important factor leading to the better predicative accuracy of the m-JIS score.

One of the differences between these scoring models is the cutoff level of tumor morphology. In the m-CLIP scoring system, tumor morphology is evaluated by whether it is uninodular and has a maximum diameter greater than 10 cm. In our patient sample, only 7.6% of patients possessed tumor(s) larger than 10 cm. As tumor size is one of the most important factors used to select treatment modalities, more precise evaluation of tumor character is useful for more accurately predicting outcome in patients with HCC. Another difference is the tumor marker included in the m-CLIP score. We previously observed that AFP is an independent prognostic factor for patients with HCC who undergo liver resection.<sup>6</sup> Protein induced by vitamin K absence or antagonist-II (PIVKA-II) has also been shown to be a useful prognostic factor in a univariate analysis.<sup>6</sup> Recently, Marrero et al.<sup>22</sup> reported that PIVKA-II is more sensitive and specific than AFP for differentiating HCC from other nonmalignant chronic liver diseases. *Leuculinaris* agglutinin-reactive AFP is reported to be another useful predictor in patients with HCC.<sup>23</sup> These tumor markers should be evaluated further to determine their suitability for use in a prognostic staging system for patients with HCC.

In conclusion, the m-JIS score can be effectively used to predict the prognosis of HCC patients in Japan, where HCC screening programs for high-risk populations have been established. Next, we plan to validate the usefulness of this scoring system in non-Japanese patient populations.

*Acknowledgments.* We thank especially all members of The Liver Cancer Study Group of Japan who completed the questionnaires.

## References

1. Bosch FX, Ribes J, Borrás J. Epidemiology of primary liver cancer. *Semin Liver Dis* 1999;19:271–85.
2. Ikai I, Arii S, Ichida T, Okita K, Omata M, Kojiro M, et al. Report of the 16th follow-up survey of primary liver cancer. *Hepatol Res* 2005;32:163–72.
3. The Liver Cancer Study Group of Japan. Primary liver cancer in Japan. Clinicopathologic features and results of surgical treatment. *Ann Surg* 1990;211:277–87.
4. The Liver Cancer Study Group of Japan. Predictive factors for long term prognosis after partial hepatectomy for patients with hepatocellular carcinoma. *Cancer* 1994;74:2772–80.
5. Arii S, Yamaoka Y, Futagawa S, Inoue K, Kobayashi K, Kojiro M, et al. Results of surgical and nonsurgical treatment for small-sized hepatocellular carcinomas: a retrospective and nationwide survey in Japan. The Liver Cancer Study Group of Japan. *Hepatology* 2000;32:1224–9.
6. Ikai I, Arii S, Kojiro M, Ichida T, Makuuchi M, Matsuyama Y, et al. Re-evaluation of prognostic factors for survival after liver resection in patients with hepatocellular carcinoma in a Japanese nationwide survey. *Cancer* 2004;101:796–802.
7. Live Cancer Study Group of Japan. General rules for the clinical and pathological study of primary liver cancer. 2nd English ed. Tokyo: Kanehara; 2003.
8. Wildi S, Pestalozzi BC, McCormack L, Clavien PA. Critical evaluation of the different staging systems for hepatocellular carcinoma. *Br J Surg* 2004;91:400–8.
9. The Cancer of the Liver Italian Program (CLIP) investigators. A new prognostic system for hepatocellular carcinoma: a retrospective study of 435 patients. *Hepatology* 1998;28:751–5.
10. Llovet JM, Bru C, Bruix J. Prognosis of hepatocellular carcinoma: the BCLC staging classification. *Semin Liver Dis* 1999;19:329–38.
11. Leung TW, Tang AM, Zee B, Lau WY, Lai PB, Leung KL, 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.
12. Kudo M, Chung H, Osaki Y. Prognostic staging system for hepatocellular carcinoma (CLIP score): its value and limitations, and a proposal for a new staging system, the Japan Integrated Staging score (JIS score). *J Gastroenterol* 2003;38:207–15.
13. Kudo M, Chung H, Haji S, Osaki Y, Oka H, Seki T, et al. Validation of a new prognostic staging system for hepatocellular carcinoma: the JIS score compared with the CLIP score. *Hepatology* 2004;40:1396–405.
14. Liang, K-Y, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986;73:13–22.
15. Okuda K, Ohtsuki T, Obata H, Tomimatsu M, Okazaki N, Hasegawa H, et al. Natural history of hepatocellular carcinoma and prognosis in relation to treatment. Study of 850 patients. *Cancer* 1985;56:918–28.
16. International Union Against Cancer. TNM classification of malignant tumor. 6th ed. Hoboken, NJ: Wiley; 2002.
17. Makuuchi M, Belghiti JB, Belli G, Fan ST, Lau JY, Ringe B, et al. IHPBA concordant classification of primary liver cancer: working group report. *J Hepatobiliary Pancreat Surg* 2003;10:26–30.
18. Lau H, Man K, Fan ST, Yu WC, Lo CM, Wong J. Evaluation of preoperative hepatic function in patients with hepatocellular carcinoma undergoing hepatectomy. *Br J Surg* 1997;84:1255–9.
19. 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. *Hepatology* 2000;31:840–5.
20. Levy I, Sherman M. Liver Cancer Study Group of the University of Toronto. Staging of hepatocellular carcinoma: assessment of the CLIP, Okuda, and Child-Pugh staging systems in a cohort of 257 patients in Toronto. *Gut* 2002;50:881–5.