

Tumor and liver determinants of prognosis in unresectable hepatocellular carcinoma: a large case cohort study

Brian I. Carr · Petr Pancoska · Robert A. Branch

Received: 20 December 2007 / Revised: 30 August 2009 / Accepted: 26 October 2009 / Published online: 24 December 2009
© Asian Pacific Association for the Study of the Liver 2009

Abstract

Objectives 967 patients with unresectable and untransplantable, biopsy-proven hepatocellular carcinoma (HCC) were prospectively evaluated at baseline and followed up till death. Survival was the end point.

Results We found that male gender, ascites, cirrhosis, portal vein thrombosis (PVT), elevated AFP or bilirubin, or alkaline phosphatase, were each statistically significant adverse prognostic factors. Patients with normal AFP survived longer than those with elevated AFP, even in the presence of PVT, large or bilobar tumors or cirrhosis. We used a bivariate analysis to separate patient sub groups based on liver function and tumor characteristics and found clear discrimination in survival between subsets; in addition both cirrhosis and presence of PVT were significant factors. We also used a purely mathematical approach to derive subgroups and a prognostic model for individual patients. Interestingly, the two approaches gave similar predictive information, which opens the possibility of a more detailed mathematical analysis in the future. The results of this large dataset show that amongst non-surgical HCC patients, there are clear subsets with longer survival.

Conclusion The data supports the concept of heterogeneity of HCC. The three factors, bilirubin, AFP, and PVT predominate in prognosis.

Abbreviations

HCC	Hepatocellular carcinoma
AFP	Alpha-fetoprotein
HCV	Hepatitis C virus
ALKP	Alkaline phosphatase
GGTP	Gamma glutamyl transpeptidase
PVT	Portal vein thrombosis
CT	Computerized axial tomography scan
HR	Hazard ratio
CI	Confidence interval

Introduction

Clinical observations of HCC across the world suggest differences in causes, such as HBV, HCV, alcoholism or aflatoxin B1 exposure [1], as well as locations [2]. However, a common major difficulty in the assessment of the natural history in many studies is the arbitrary time at which diagnosis is made in each patient. This includes early identification due to screening, or to diagnosis later in the disease progression consequent on complications from tumor growth, or from hepatic failure. Many scoring and classification systems have taken the twin issues of cirrhosis and tumor extent into account, including those of Okuda, CLIP, BCLC, CUPI, SLiDE, JIS, and Tokyo scores, amongst others [3–11] and several papers have recently compared multiple staging systems [12–23]. They have been mainly intended to aid in patient selection for surgery, loco-regional therapy such as RFA or TACE, or palliative care. However, in the US, where there is no national screening program, the majority of HCC patients present at an advanced phase of their disease, either due to the severity of the underlying cirrhosis or the extent of their

B. I. Carr (✉)
Liver Tumor Program, Kimmel Cancer Center,
Thomas Jefferson University, Bluemle Building, Room 1002,
233 S. 10th Street, Philadelphia 19107, PA, USA
e-mail: brian.carr@jefferson.edu;
brian.carr@kimmelcancercenter.org

P. Pancoska · R. A. Branch
Center for Clinical Pharmacology, University of Pittsburgh,
Philadelphia, USA

HCC. There has not so far, to our knowledge, been a large prospective study of these advanced unresectable HCC patients; although, several small series having evaluated prognostic factors in these patients [8, 20, 24–29].

The difficulty of analysis in HCC is compounded by the fact that tumor growth can cause a variety of complications due to direct tumor extension and a reduction in residual hepatic function can result from either the associated cirrhosis or from the tumor extension, so that assumptions based on independent variables such as conventional regression analysis may not be valid. An alternative strategy to address the problem of variation in the time of diagnosis is to use a subset analysis strategy. In this approach, the major variables relating to the stage in the natural history are taken into consideration to identify more homogenous and comparable populations at the time of clinical presentation. The impact of the variable of interest on clinical outcome is then compared in each patient subset. This approach requires a substantial sample size and uniformity in both data collection as well as treatment approach.

The recent availability of information obtained at baseline assessment with prolonged follow up of clinical outcomes from an exceptionally large case cohort of patients provides an opportunity to use this strategy. All the patients were referred to a single physician (BIC) at one institution over 17 years for medical management of unresectable HCC. We have addressed the hypothesis that HCC represents heterogeneous disease groups.

If valid, the clinical implication for this hypothesis is that patients with identifiable profiles or phenotypes at disease presentation could be anticipated to have differences in prognosis. We have also derived a potentially useful clinical scoring system from a purely data-derived approach that was harmonized with the physician-driven intuitive strategy, to predict the prognosis of an individual case and contrast the utility of the two approaches.

Methods

Clinical methods

This prospective 17-year experience was acquired in HCC patients being seen by a single liver oncologist. 967 patients with HCC that were considered unsuitable for resection or liver transplantation were managed with TACE. On initial clinical evaluation, all patients had baseline CBC, liver function tests, alpha-fetoprotein (AFP), hepatitis serology, serum creatinine, physical examination, liver and tumor biopsy, and a triphasic helical computed axial tomography (CT) scan of the chest, abdomen and

pelvis. Liver and tumor biopsies were obtained. The data and CT descriptors were entered into an HCC database intended for follow-up and analysis. All patients were followed for the time of their survival either through this clinic, or through close liaison through their primary care physician.

Statistical methods

Overall survival was the only end point used, defined as the time between the date of diagnosis and date of death. In order to devise a potentially useful clinical scoring system from this large dataset that could be applied to the prognosis of an individual patient, we undertook a detailed inspection of the plots of survival versus the laboratory values of bilirubin, alkaline phosphatase, SGOT, and GGTP. As the relationship of each parameter levels to survival had a non-linear correlation, we partitioned the data of each parameter into multiple range sets to better capture this reality in the scoring. We converted the recorded lab values to the same relative scale (0–100%) for each parameter, to permit comparison of value histograms. We found the overall distributions of all these parameter value frequencies were the sum of five components, expressed as five Gaussian distribution curves. Each of these ranges was characterized by the most frequently observed value (maxima of components). We calculated the average characteristic concentration value for each of the five Gaussian distribution ranges and converted them into relative numbers (input) with respect to the maximum for each parameter. The clinical descriptive parameters (cirrhosis, ascites, PVT, logAFP) were converted into a simple numerical scale and combined with the transformed characteristic laboratory variables input. A set of this combined data for each individual patient's unique descriptors was used to characterize each patient and to derive appropriate weighting of each variable in the scoring system. Logistic regression analysis was then performed to predict a short- or long-survival category for each patient. To identify the best threshold between short- and long-survival, we repeated the regression analyses for several thresholds between 30 and 180 days and found the optimal prediction was for a threshold of 90 days. In the final step, we derived clinically usable scores by multiplying the characteristic inputs by their regression coefficients from the logistic regression model, to identify a variable-specific contribution to identify a single variable (sum) that if negative (<0), indicates a survival longer than 90 days; and if positive (>0), indicates a survival shorter than 90 days. This scoring system is entirely data-driven, rather than hypothesis-driven.

Results

Characterization of the heterogeneity of the patient data

This is data from a practice, not a clinical trial. In the optimal design of a gold-standard clinical trial, the study design minimizes the impact of many complex aspects of patient disease, while these aspects have clear impact on the information in the practice-based clinical data. This feature of the clinical practice databases requires using special methods for their analysis. In this article, we show that in the well-curated and ideally recorded data from normal clinical practice for cohorts of thousands of patients, the information about many aspects of the disease/liver cancer/HCC heterogeneity are indeed present and can be extracted.

In our approach to extracting the clinically relevant information for patients with HCC, we took advantage of our understanding of the important factors that characterize the uniqueness of information in the clinical practice data. Figure 1 summarizes schematically our understanding as interplay of the four factors: data heterogeneity, stochasticity, ambiguity, and coherence.

These terms summarize in a compact way the following features of the information in the clinical practice data: 1. stochastic factor is defined by the uncertainty of the actual disease status of patients that come for treatment in random stages of the disease. 2. Factor of heterogeneity (the one we analyze in more detail in this article) represents the possibility of distinct underlying biology or pathophysiology that might be grouped together under the umbrella of a common diagnosis. 3. Factor of ambiguity represents the possibility that the same values of one clinical parameter can have different significance for patients with different clinical or pathophysiological status. 4. Factor of coherence represents the common sense clinician's understanding that if certain combination of one series of clinical parameters is found significantly more often (=coherently) than other combinations, then such an observation might have profound clinical relevance. Figure 1 emphasizes the fact that these four factors cannot be treated independently, because the actual status of one influences the clinical impact of all

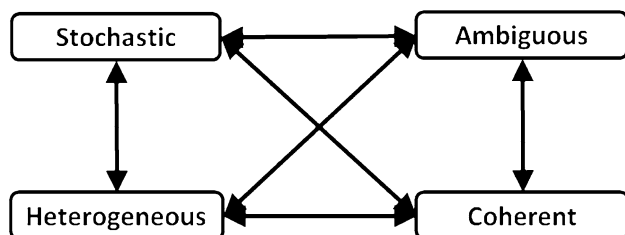


Fig. 1 The four general factors describing the interplay of information categories in a clinical practice database

the others. Thus, different combinations of these four factors provide different angles through which we can see, process, and analyze informative structures of the raw clinical data (Fig. 1). When used in combination to design the mathematics of the data analysis, optimal extraction of the relevant information from normal clinical data can be achieved.

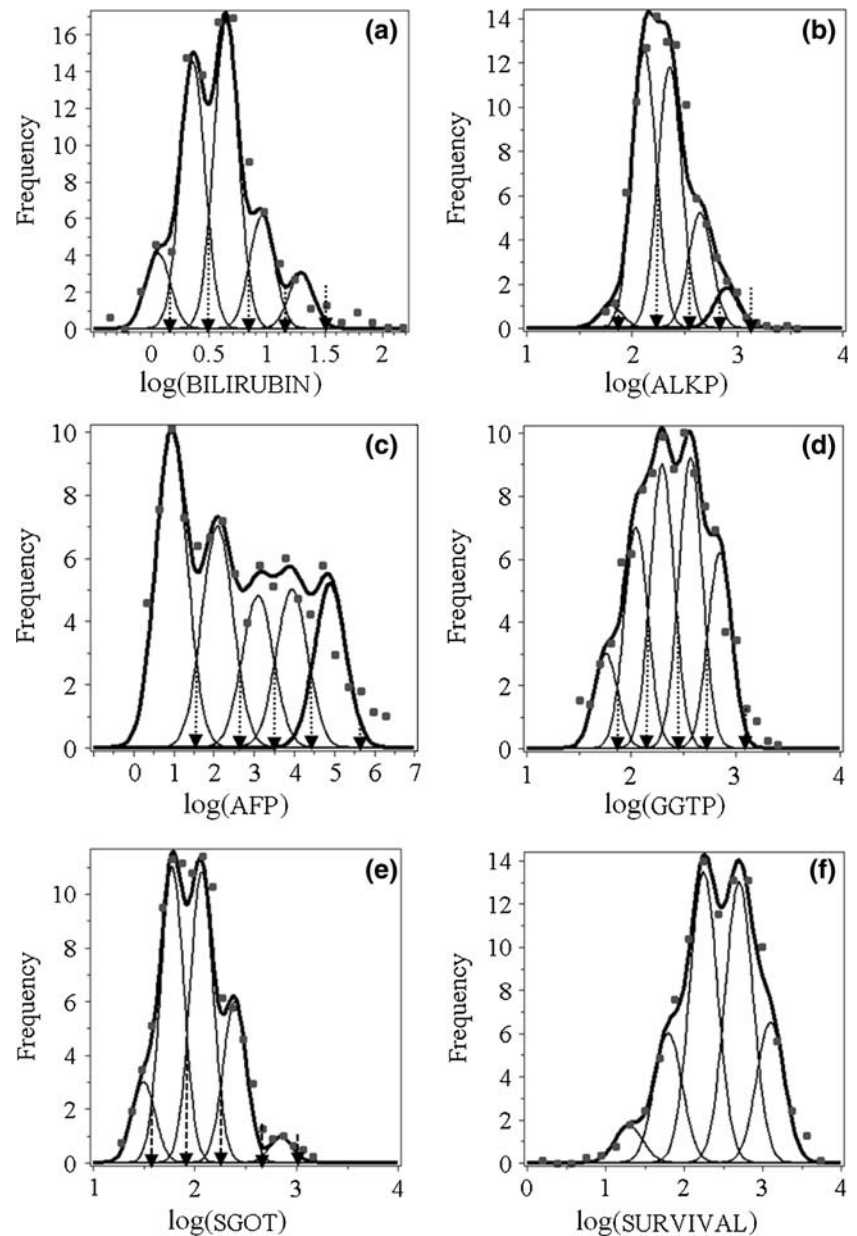
In this article, we approached the analysis through the perspective of a combination of heterogeneity and coherence of the information in the HCC dataset. We first asked whether there was evidence of heterogeneity of the disease in the HCC dataset. To approach this, we took advantage of the large number of patients in this database, which facilitates our construction of histograms of the clinical parameter values that represented the probability distributions of these values. We next examined whether the distributions of the values of these parameters were multi-component in a way that will provide significant information about subsets of patients who carry the respective HCC disease subtypes.

Mathematical approach to the heterogeneity and coherence

We initially computed the histograms of the frequencies of levels of our selected clinical parameters. Inspection of these histograms (Fig. 2) showed that they are indeed multi-component. Next, we statistically tested the 'fit' of all histograms by a single Gaussian component. In all cases, the fit failed the normality and constant variance tests (these tests indicated with $P < 0.001$ that hypothesized that the parameter value histograms are single Gaussian distributions is false). We also recorded the regression coefficients from these single-component fits, to compare them to the regression coefficients from the multi-component fits described below. In all cases, the multi-component fit resulted in significant improvement of the regression ($P < 0.01$) and resulted also in passing the constant variance tests (which mean that the difference between the reconstructed and the actual histogram does not contain systematic deviations from zero).

The key issue for a mathematical approach to the quantitative description of the multiple components in the HCC clinical data histograms is to select the optimal number of components that combine into the overall observed distribution. We used an iterative procedure, in which we applied non-linear least square fitting of the actual histogram by the sum of the variable number of normal distributions that sum into the actual histogram of parameter levels for each parameter. We were seeking the optimal number of sub-distributions, which we define as the minimal number of the component normal distributions that is needed to fit the shapes of the histograms of all

Fig. 2 Histogram depicting the heterogeneity in the HCC clinical data



parameters. From the clinician's perspective, this feature of our mathematical algorithm represents the quantification of the common sense clinician's understanding that the disease is heterogeneous. This heterogeneity is more evident if it is found in all the parameters that are being studied and if mathematical treatment of each parameter is coherent with the treatment of all other parameters.

Figures 2 and 3 show the optimal result by repeating the fitting of all parameter histograms, we found that we could describe all these histograms with only five-component normal distributions. This indicates that there is not a limitless heterogeneity of HCC clinical data, but rather that the data for each parameter can be described in terms of only five groups or components. Thus, the data show us

that we can treat the heterogeneity of information in HCC databases by dividing the data into six intervals, which describe internally homogeneous patient subsets.

This required a definition of five threshold values for each parameter. We indicate by the dotted arrows in Fig. 2 how we used the five-component decompositions of parameter histograms for this purpose. Each of the five-parameter ranges shown in Table 1 were derived from the points at which the component Gaussian distributions crossed over each other in Fig. 2. Our results indicate that while in the clinical practice data the different subgroups cannot be completely separated, our data-driven approach to describe these five thresholds of each parameter is the best choice that maximizes the separation of the coherent

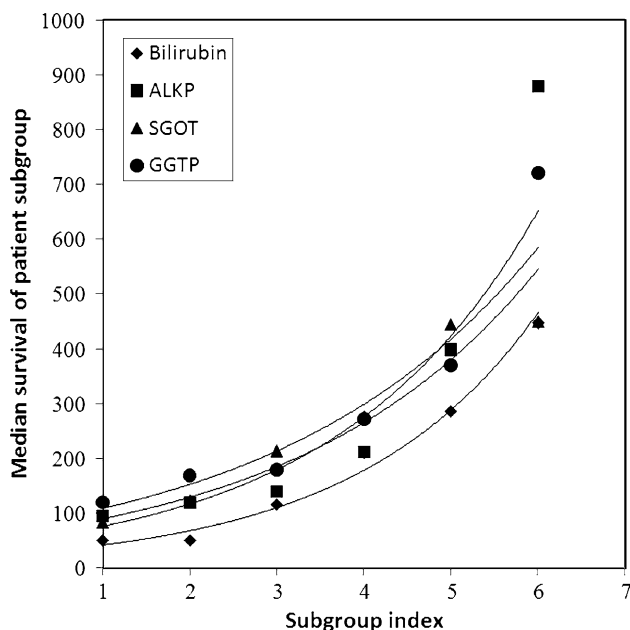


Fig. 3 Relationship of range 1–6 of the serum-based clinical parameters as derived by our data-driven definition of threshold values (see Table 1). Thus, index 1 for bilirubin represents values above 9.5, index 2 bilirubin values 4.5–9.5, index 3 range 2.0–4.4, index 4 range 1.0–1.9, index 5 range 0.5–0.9, and index 6 range 0–0.4 mg/dl. For ALKP, index 1 corresponds to range above 900, index 2 to range 435–900 U/100 ml etc. also for other parameters as shown in Table 1. The lines show that relationship between these range characteristics and median survival of all patients with parameter values in the respective ranges can be described as exponential, with statistical significance ($R^2 > 9.4$ and $P < 0.0001$)

levels of parameter ranges. It is also clear that any other arbitrary way of choosing clinical ranges (for example dividing the parameter values into equal intervals) would result in suboptimal result, where the data from different categories are more mixed when compared to our optimal selection, which reflect best the actual structure of the heterogeneity of HCC data. This is the unique feature of our clinical data that we were able to quantitatively capture by our analysis method, which in turn facilitates our treatment of the data and its application to clinical prognosis. Thus, Table 1 ranges were determined as is shown by arrows in Fig. 2 for each parameter (ALKP, AFP, GGTP, SGOT).

Validation through reconstruction of short- and long-survival categories

While the coherence factor of a common number of components in the distributions of clinical parameter data helps to extract the relevant information about the heterogeneity of the ranges of these clinical parameters, we still needed to validate the clinical relevance of this heterogeneity. First, we show in Table 1 that there is a strong correlation

between the ranges 1–6 of levels of all relevant clinical parameters and median survival, as calculated for patients with the levels of individual parameters within each interval 1–6.

We used this as indication that it is possible to use the individual laboratory result to assign any patient to one of the six categories and then to develop the classification scheme allowing to characterize the patient as short- or long-surviving. To this purpose, we used the logistic regression algorithm with variable selection to determine the optimal combination of the serum-based and pathophysiological parameters. Moreover, we again adopted the iterative optimization approach to derive data-driven rather than arbitrary definition of the short- and long-survival. As is described in “Methods”, we obtained the best classification performance of our logistic regression algorithm for setting threshold between the short- and long-survival at 90 days. With this threshold, we were able to develop a logistic discrimination model that in 10-fold cross-validation classified correctly 85% cases (ROC area 0.79). For convenience and application in clinical practice, we converted the optimized result into the simple prognostic scheme, which is described in the next paragraph.

A data-driven analysis approach to survival prediction

In order to complement the clinically intuitive sub grouping strategy above, with a purely data-driven strategy, we used a stochastic approach (described in the “Methods”) to identify six characteristic ranges of serum bilirubin, alkaline phosphatase, SGOT, and GGTP in combination with the presence or absence of cirrhosis, ascites, PVT, or abnormal AFP to derive scoring values of each variable in a scheme to predict long- or short-outcome of survival. Optimization procedures identified 90 days of survival as providing the greatest discrimination. Using dichotomization at this time point, scores for each variable are presented in Table 1. We illustrate contrasting applications of this model, in two different patients. At baseline, patient A had no PVT (code 0), with cirrhosis (code 1) with ascites (code 1), baseline AFP of 537 ng/ml, bilirubin 2.0 mg/dl, alkaline phosphatase 155 U/100 ml, SGOT 168 U/100 ml, and GGTP 121 U/100 ml. Using Table 1, the scoring for this patient is PVT score 0, cirrhosis score -40 , ascites score 75, AFP score 56, bilirubin score -235 , alkaline phosphatase score -115 , SGOT score -35 , and GGTP score 15. The HCC score calculation for this patient is $290 + 0 - 40 + 75 + 56 - 235 - 115 - 35 + 15 = 11$. As this result is positive, the scoring predicts class “A”, survival, with prognosis shorter than 90 days. In contrast, at baseline evaluation, patient B had PVT only on the left side (code 1), with cirrhosis (code 1), without ascites (code 0), baseline levels of AFP 536 ng/ml, bilirubin 0.3 mg/dl, alkaline phosphatase

Table 1 Data-driven scoring system for survival prognosis

Parameter	Score	Median survival (days)	Confidence interval (\pm days)
Base value	290		
Bilirubin range			
Above 9.5	-100	50	54
9.4–4.5	-180	50	35
4.4–2.0	-235	116	73
1.9–1	-300	210	57
0.9–0.5	-370	286	68
0.4–0	-460	447	107
ALKP range			
Above 900	-40	96	133
899–435	-67	120	144
434–240	-90	140	50
239–125	-115	212	46
124–55	-140	397	85
54–0	-175	880	675
SGOT range			
Above 290	-20	83	79
289–140	-35	122	54
139–80	-50	213	67
79–40	-60	275	53
39–15	-75	445	150
14–0	-90	>450	150
GGTP range			
Above 770	5	120	67
769–390	9	169	62
389–180	12	180	49
179–80	15	272	84
79–35	19	370	106
34–0	25	721	408
PVT code^a			
-1	15	102	47
0	0	313	49
1	-15	167	86
2	-30	163	72
3	-45	125	117
Cirrhosis code^b			
-1	40	121	121
0	0	343	105
1	-40	193	33
Ascites code^c			
-1	-75	141	109
0	0	341	49
1	75	117	48
AFP			
	$20 \times \log(\text{AFP})$		
	<25	340	71
	25–2000	203	93
	>2000	148	38

^a Codes: -1 = not determined, 0 = present, 1 = portal branch only, 2 = main portal vein, 3 = main plus branch portal vein

^{b,c} Cirrhosis and ascites codes: -1 = not determined, 0 = absent, 1 = present

263 U/100 ml, SGOT 83 U/100 ml, and GGTP 98 U/100 ml. Applying this information to Table 1, the PVT score is -15 , cirrhosis score -40 , ascites score 0 , AFP score 55 , bilirubin score -460 , alkaline phosphatase score -90 , SGOT score -50 , and GGTP score 15 . The HCC score calculation for this patient is $290 - 15 - 40 + 0 + 55 - 460 - 90 - 50 + 15 = -295$. As this result is negative, the scoring predicts class “B”, survival prognosis of longer than 90 days.

In a 10-fold cross-validation, this model correctly predicted survival in 80% of cases with a true positive rate of 94% for long survival and 36% for short survival. To verify the statistical significance of our model, we also performed a 10,000-fold permutation of the short- and long-survival prognoses and tested the likelihood of obtaining the result of the model purely by chance. In none of these multiple random permutation tests, were results comparable (equal) or better than the model obtained either for the short- or for the long-survival class. Therefore, the significance of our results in terms of P value is better than 0.0001 and that this quantitative data-driven scoring system has high degree of predictability.

For direct comparison of two approaches, namely a clinical intuition-driven when compared with a data-driven partitioning of patients into clinically relevant groups, we calculated median survival and 95% confidence interval for each sub-group of patients with their parameters within the ranges shown in Table 1. We found that value $1.2 \times [(Average\ of\ all\ means\ in\ Table\ 1) - 115]$ correlates best with the actual survival time (data not shown). For our two example patients these prognostic calculations are as follows:

$$\begin{aligned} \text{Patient A prognosis} &= 1.2 \times [(1/8.0) * (313 + 193 + 117 \\ &\quad + 148 + 116 + 212 + 122 + 272) \\ &\quad - 115] \\ &= 86 \text{ days.} \end{aligned}$$

$$\begin{aligned} \text{Patient B prognosis} &= 1.2 \times [(1/8.0) * (162 + 193 + 341 \\ &\quad + 203 + 447 + 140 + 213 + 272) \\ &\quad - 115] \\ &= 158 \text{ days} \end{aligned}$$

This is in reasonable agreement with the predicted survival of 90–180 days and 150–270 days from Fig. 1.

Discussion

The present series is the largest single site report of consecutive patients with unresectable HCC in the U.S.A lengthy follow up of time from diagnosis to death in this unique cohort provides insight into the pathophysiology of

HCC progression, and suggests that there is heterogeneity of disease states that are included within this diagnosis to influence rate of disease progression. Overall, once HCC is not resectable or transplantable, the prognosis at time of clinical evaluation is limited ($9 \pm$ months, median \pm 95% CI). However, within this cohort, there are features that determine whether this prognosis is restricted to terminal patient care (<3 months), or those whose life can be expected to extend for longer. We suggest that this discrimination is of clinical value in selecting between therapeutic choices of terminal palliative care, or an intervention designed to both ameliorate symptoms and possibly prolong survival.

HCC is an unusual cancer in that its progression can be detrimental to the patient due to two closely intertwined but discrete processes, namely tumor growth and decreased residual hepatic function, each of which can be evaluated from a variety of indices and each of which can independently be the cause of death. This was recognized in one of the earliest classification systems [3]. It is also evident that the time of diagnosis within the disease progression in any individual patient is arbitrary. Thus, an overall analysis of individual factors influencing survival is complex and confounded by interdependency between variables. The size of this patient cohort has permitted an alternative to the usual analysis strategy, by creating mutually exclusive categorizations for each patient into subsets based on CT scan and biochemical variables obtained at the time of diagnosis and comparing times of survival between subsets. We have followed a progressive strategy. We have taken variables selected to represent characteristics of the tumor (number, size, evidence of extension as judged by portal vein thrombosis) and residual hepatic function (liver function tests). We infer that HCC in the absence of cirrhosis represents a different pathophysiological entity with different rate of tumor progression than HCC in the presence of cirrhosis.

Using a similar strategy and the same logic, we also suggest that the presence of AFP above the normal level for our lab, is a similar negative prognostic indicator of survival in each of the patient subgroups, when defined by liver function and tumor extension, as previously described [30, 31]. Thus, a similar inference can be made, namely that AFP positive and AFP negative HCC represent discrete phenotypes, with different rates of tumor progression. Most classification systems have also recognized these same variables as being predictors of long-term outcome. This analysis differs, in that it is confined only to patients who were considered untreatable and deemed to be unresectable and untransplantable (or untreatable by any local ablative technique RFA, PEI), as well as the large size of this cohort study in the US. It also attempts to deal with the multiple factors involved (such as cirrhosis, ascites, liver

function tests), which not being independent variables, require a slightly different approach to analysis.

Several new HCC tumor markers have recently been clinically evaluated and found to be useful in HCC diagnosis, especially AFP-L3 and DCP [29, 32–36] as well as possibly hsp27, osteopontin, and glypican 3 [37–41]. The rapidly evolving fields of genomics and proteomics are now finding HCC-associated gene products that will likely find future usefulness in further defining HCC prognostic subsets [42–44], including metastasis potential [45–47].

Clinical anecdotal case reporting is subject to a variety of biases. The problems that are inherent in cohort analysis are present here, but have been minimized to the extent possible. At the outset of creating a specialized clinical HCC program, a prospective study design was developed, of creating a phenotypic description of each patient with common historical, clinical, CT scan imaging, tumor marker, and biochemical measures. Each consecutive patient referred to a single clinician has been included, and most importantly, follow up to death from cancer is almost complete, due to a prospective and involved commitment of the physician and nurses to patient care and follow-up. This has been sustained over the 17 years of the clinical service. The regimen of cisplatin-TACE was used throughout and the patient cohort was homogeneous in that none were subjected to subsequent surgery. This follow-up information provides the very hard data end point of death.

There are, however limitations. The major bias arises from the success in the clinic in attracting patients from a wide geographic location based on physician referral advice. The treatments of cisplatin-TACE or transhepatic arterial chemo-embolization which were offered to these patients have only a small proven benefit on survival [47]. This is also a strength of this study, as TACE has then become a surrogate for studying contemporary natural history. Thirdly, quantitative approaches to understanding the pathophysiology of HCC growth and loss of residual liver function are in their infancy, and the measures used are known to be insensitive. Fourthly, any inferences drawn from this study have to be limited to the US population of HCC patients.

We have used a novel data-driven strategy that takes away arbitrary dichotomized decisions in analysis, to create a multivariate analysis in which each of the variables was normalized and contrasted against the others, to provide relative scoring. This was then used to dichotomize survival and identify longer or shorter survival relative to 90 days. We were surprised at the concordance of the continuous variable datasets to each identify six sub-populations and impressed by the ability and statistical significance to predict the outcome. We therefore consider that the relatively simple procedure presented in Table 1 augments other standard simpler classification systems.

Thus, clinical selection for alternative therapeutic guidelines can be aided by a combination of two simple numerical algorithms.

Whether this extensive experience is considered from the perspective of a clinical algorithm of a tree-based subgroup analysis as is usual, or from a data-driven multivariate analysis (Table 1), it is apparent that the major predictor of a poor outcome is the serum bilirubin at the time of diagnosis. This observation implies that maintenance of residual liver function is more important than features attributed to tumor growth and extension. If anything, a previous subgroup analysis of the present series which estimated that 40% of HCC patients died from liver failure rather than from tumor progression [29], was an under-estimate. This reinforces the dual complexity of changes in liver function and HCC tumor growth.

In summary, by the time HCC has progressed to be not amenable to resection or liver transplantation, the prognosis is limited. There are, however, discrete characteristics that can be shown to influence survival. Of these, the three most dominant are PVT, elevated AFP, and loss of residual hepatic function (elevated bilirubin levels). Although several classification systems exist, these three factors predominate in prognostic significance for unresectable HCC. It is also reassuring that clinical intuition based on only three easily remembered variables is equivalent, from the perspective of clinical decision-making, to a rigorous, purely data-derived strategy of variable weighting. We also suggest that a bilirubin level >1.5 mg/dl and/or presence of PVT and elevated AFP should be taken into account in future trials of innovative therapies, to provide a strategy for stratification.

References

1. El-Serag KR. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. *Gastroenterology* 2007;132:2557–2576
2. Cho SJ, Yoon JH, Hwang SS, et al. Do young hepatocellular carcinoma patients with relatively good liver function have poorer outcomes than elderly patients? *J Gastroenterol Hepatol* 2007;22:1226–1231
3. Okuda K, Ohtsuki T, Obata H, et al. Natural history of hepatocellular carcinoma and prognosis in relation to treatment. *Cancer* 1985;56:918–928
4. Cancer of the Liver Italian Program (CLIP) investigators. A new prognostic system for hepatocellular carcinoma: a retrospective study of 435 patients. *Hepatology* 1998;28:751–755
5. 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
6. Llovet JM, Fuster J, Bruix J. The Barcelona approach: diagnosis, staging, and treatment of hepatocellular carcinoma. *Liver Transpl* 2004;10:S115–S120

7. Llovet JM, Bru C, Bruix J. Prognosis of hepatocellular carcinoma: the BCLC staging classification. *Semin Liver Dis* 1999;19:329–338
8. Tateishi R, Yoshida H, Shiina S, et al. Proposal of a new prognostic model for hepatocellular carcinoma: an analysis of 403 patients. *Gut* 2005;54:419–425
9. Chevret S, Trinchet JC, Mathieu D, et al. A new prognostic classification for predicting survival in patients with hepatocellular carcinoma. *J Hepatol* 1999;31:133–141
10. Leung TW, Tang AM, Zee B, et al. Construction of the Chinese University Prognostic Index for hepatocellular carcinoma and comparison with the TNM staging system, the Okuda staging system, and the Cancer of the Liver Italian Program staging system: a study based on 926 patients. *Cancer* 2002;94:1760–1769
11. Omagari K, Honda S, Kadokawa Y, et al. Preliminary analysis of a newly proposed prognostic scoring system (SLiDe score) for hepatocellular carcinoma. *J Gastroenterol Hepatol* 2004;19:805–811
12. Levy I, Sherman M. Staging of hepatocellular carcinoma: assessment of the CLIP, Okuda, and Child-Pugh staging systems in a cohort of 257 patients in Toronto. *Gut* 2002;50:881–885.
13. Giannini E, Rizzo D, Botta F, et al. Prognosis of hepatocellular carcinoma in anti-HCV positive cirrhotic patients: a single-centre comparison amongst four different staging systems. *J Intern Med* 2004;255:399–408
14. Marrero JA, Fontana RJ, Barrat A, et al. Prognosis of hepatocellular carcinoma: comparison of 7 staging systems in an American cohort. *Hepatology* 2005;41:707–716
15. Huang YH, Chen CH, Chang TT, et al. Evaluation of predictive value of CLIP, Okuda, TNM and JIS staging systems for hepatocellular carcinoma patients undergoing surgery. *J Gastroenterol Hepatol* 2005;20:765–771
16. Grieco A, Pompili M, Caminiti G, et al. Prognostic factors for survival in patients with early-intermediate hepatocellular carcinoma undergoing non-surgical therapy: comparison of Okuda, CLIP, and BCLC staging systems in a single Italian centre. *Gut* 2005;54:411–418
17. Pascual S, Zapater P, Such J, et al. Comparison of staging systems to predict survival in hepatocellular carcinoma. *Liver Int* 2006;26:673–679
18. Kondo K, Chijiwa K, Nagano M, et al. Comparison of seven prognostic staging systems in patients who undergo hepatectomy for hepatocellular carcinoma. *Hepatogastroenterology* 2007;54:1534–1538
19. Huo TI, Lin HC, Hsia CY, et al. The model for end-stage liver disease based cancer staging systems are better prognostic models for hepatocellular carcinoma: a prospective sequential survey. *Am J Gastroenterol* 2007;102:1920–1930
20. Georgiades CS, Liapi E, Frangakis C, et al. Prognostic accuracy of 12 liver staging systems in patients with unresectable hepatocellular carcinoma treated with transarterial chemoembolization. *J Vasc Interv Radiol* 2006;17:1619–1624
21. Minagawa M, Ikai I, Matsuyama Y, et al. Staging of hepatocellular carcinoma: assessment of the Japanese TNM and AJCC/UICC TNM systems in a cohort of 13, 772 patients in Japan. *Ann Surg* 2007;245:909–922
22. Chung H, Kudo M, Takahashi S, et al. Review of current staging systems for hepatocellular carcinoma. *Hepatol Res* 2007;37:S210–S215
23. Guglielmi A, Ruzzenente A, Pachera S, et al. Comparison of seven staging systems in cirrhotic patients with hepatocellular carcinoma in a cohort of patients who underwent radiofrequency ablation with complete response. *Am J Gastroenterol* 2008;103(3):597–604
24. Pawarode A, Voravud N, Sriuranpong V, et al. Natural history of untreated primary hepatocellular carcinoma: a retrospective study of 157 patients. *Am J Clin Oncol* 1998;21:386–391
25. Llado L, Virgili J, Figueras J, et al. A prognostic index of the survival of patients with unresectable hepatocellular carcinoma after transcatheter arterial chemoembolization. *Cancer* 2000;88:50–57
26. O'Suilleabhain CB, Poon RT, Yong JL, et al. Factors predictive of 5-year survival after transarterial chemoembolization for inoperable hepatocellular carcinoma. *Br J Surg* 2003;90:325–331
27. Yamasaki T, Kimura T, Kurokawa F, et al. Prognostic factors in patients with advanced hepatocellular carcinoma receiving hepatic arterial infusion chemotherapy. *J Gastroenterol* 2005;40:70–78
28. Kogo M, Kano A, Kiuchi Y, Mitamura K, Yoneyama K. Prognostic index for survival in patients after treatment for primary hepatocellular carcinoma. *Dig Dis Sci* 2007;52(9):2444–2451
29. Couto OF, Dvorchik I, Carr BI. Causes of death in patients with unresectable hepatocellular carcinoma. *Dig Dis Sci* 2007;52:3285–3289
30. Dvorchik I, Carr BI. A simple prognostic scoring system for patients with unresectable hepatocellular carcinoma treated by chemo-embolization. *Cancer Detect Prev* 2007;31:154–160
31. Johnson PJ, Melia WM, Palmer MK, et al. Relationship between serum alpha-fetoprotein, cirrhosis and survival in hepatocellular carcinoma. *Br J Cancer* 1981;44:502–505
32. Wang CS, Lin CL, Lee HC, et al. Usefulness of serum des-gamma-carboxy prothrombin in detection of hepatocellular carcinoma. *World J Gastroenterol* 2005;11:6115–6119
33. Carr BI, Kanke F, Wise M, et al. Clinical evaluation of lens culinaris agglutinin-reactive alpha-fetoprotein and des-gamma-carboxy prothrombin in histologically proven hepatocellular carcinoma in the United States. *Dig Dis Sci* 2007;52:776–782
34. Toyoda H, Kumada T, Osaki Y, et al. Role of tumor markers in assessment of tumor progression and prediction of outcomes in patients with hepatocellular carcinoma. *Hepatol Res* 2007;37:S166–S171
35. Volk ML, Hernandez JC, Su GL, et al. Risk factors for hepatocellular carcinoma may impair the performance of biomarkers: a comparison of AFP, DCP, and AFP-L3. *Cancer Biomark* 2007;3:79–87
36. Shirabe K, Itoh S, Yoshizumi T, et al. The predictors of microvascular invasion in candidates for liver transplantation with hepatocellular carcinoma-with special reference to the serum levels of des-gamma-carboxy prothrombin. *J Surg Oncol* 2007;95:235–240
37. Capurro M, Wanless IR, Sherman M, et al. Glypican-3: a novel serum and histochemical marker for hepatocellular carcinoma. *Hepatogastroenterology* 2003;125:89–97
38. Feng JT, Liu YK, Song HY, et al. Heat-shock protein 27: a potential biomarker for hepatocellular carcinoma identified by serum proteome analysis. *Proteomics* 2005;5:4581–4588
39. Di Tommaso L, Franchi G, Park YN, et al. Diagnostic value of HSP70, glypican 3, and glutamine synthetase in hepatocellular nodules in cirrhosis. *Hepatology* 2007;45:725–734
40. Kim J, Ki SS, Lee SD, et al. Elevated plasma osteopontin levels in patients with hepatocellular carcinoma. *Am J Gastroenterol* 2006;101:2051–2059
41. Lee JS, Chu IS, Heo J, et al. Classification and prediction of survival in hepatocellular carcinoma by gene expression profiling. *Hepatology* 2004;40:667–676
42. Sun W, Xing B, Sun Y, et al. Proteome analysis of hepatocellular carcinoma by two-dimensional difference gel electrophoresis: novel protein markers in hepatocellular carcinoma tissues. *Mol Cell Proteomics* 2007;6:1798–1808
43. Jia HL, Ye QH, Qin LX, et al. Gene expression profiling reveals potential biomarkers of human hepatocellular carcinoma. *Clin Cancer Res* 2007;13:1133–1139
44. Poon TC, Yip TT, Chan AT, et al. Comprehensive proteomic profiling identifies serum proteomic signatures for detection of

- hepatocellular carcinoma and its subtypes. *Clin Chem* 2003;49:752–760
45. Budhu AS, Zipser B, Forgues M, et al. The molecular signature of metastases of human hepatocellular carcinoma. *Oncology* 2005;69(Suppl 1):23–27
46. Kaposi-Novak P, Lee JS, Gomez-Quiroz L, et al. Met-regulated expression signature defines a subset of human hepatocellular carcinomas with poor prognosis and aggressive phenotype. *J Clin Invest* 2006;116:1582–1595
47. Lo CM, Ngan H, Tso WK, et al. Randomized controlled trial of transarterial lipiodol chemoembolization for unresectable hepatocellular carcinoma. *Hepatology* 2002;35:1164–1171