A Nomogram to Predict Long-term Survival After Resection for Intrahepatic Cholangiocarcinoma
An Eastern and Western Experience

Omar Hyder, MD; Hugo Marques, MD; Carlo Pulitano, MD; J. Wallis Marsh, MD; Sorin Alexandrescu, MD; Todd W. Bauer, MD; T. Clark Gamblin, MD, MS; Georgios C. Sotiropoulos, MD, PhD; Andreas Paul, MD; Eduardo Barroso, MD; Bryan M. Clary, MD; Luca Aldrighetti, MD; Cristina R. Ferrone, MD; Andrew X. Zhu, MD; Irinel Popescu, MD; Jean-Francois Gigot, MD, PhD; Gilles Mentha, MD; Shen Feng, MD; Timothy M. Pawlik, MD, MPH, PhD

IMPORTANCE Intrahepatic cholangiocarcinoma (ICC) is a primary cancer of the liver that is increasing in incidence, and the prognostic factors associated with outcome after surgery remain poorly defined.

OBJECTIVE To combine clinicopathologic variables associated with overall survival after resection of ICC into a prediction nomogram.

DESIGN, SETTING, AND PARTICIPANTS We performed an international multicenter study of 514 patients who underwent resection for ICC at 13 major hepatobiliary centers in the United States, Europe, and Asia from May 1, 1990, through December 31, 2011. Multivariate Cox proportional hazards regression modeling with backward selection using the Akaike information criteria was used to select variables for construction of the nomogram. Discrimination and calibration were performed using Kaplan-Meier curves and calibration plots.

INTERVENTIONS Surgical resection of ICC at a participating hospital.

MAIN OUTCOMES AND MEASURES Long-term survival, effect of potential prognostic factors, and performance of proposed nomogram.

RESULTS Median patient age was 59.2 years, and 53.1% of the patients were male. Most patients (74.7%) had a solitary tumor, and median tumor size was 6.0 cm. Patients were treated with an extended hepatectomy (202 [39.3%]), a hemihepatectomy (180 [35.0%]), or a minor liver resection (<3 segments) (132 [25.7%]). Most patients underwent R0 resection (87.9%), and 35.7% of patients had N1 disease. Using the backward selection of clinically relevant variables, we found that age at diagnosis (hazard ratio [HR], 1.31; \( P < .001 \)), tumor size (HR, 1.50; \( P < .001 \)), multiple tumors (HR, 1.58; \( P < .001 \)), cirrhosis (HR, 1.51; \( P = .08 \)), lymph node metastasis (HR, 1.78; \( P = .01 \)), and macrovascular invasion (HR, 2.10; \( P < .001 \)) were selected as factors predictive of survival. On the basis of these factors, a nomogram was created to predict survival of ICC after resection. Discrimination using Kaplan-Meier curves, calibration curves, and bootstrap cross-validation revealed good predictive abilities (C index, 0.692).

CONCLUSIONS AND RELEVANCE On the basis of an Eastern and Western experience, a nomogram was developed to predict overall survival after resection for ICC. Validation revealed good discrimination and calibration, suggesting clinical utility to improve individualized predictions of survival for patients undergoing resection of ICC.
Intrahepatic cholangiocarcinoma (ICC) is the second most common primary liver malignant tumor after hepatocellular carcinoma (HCC). It represents 10% to 20% of all primary liver malignant tumors, with an estimated 3100 new cases of ICC occurring every year in the United States. A relatively uncommon disease with a poor prognosis for survival, the incidence of ICC has been increasing during the last 3 decades. Surgical resection is the mainstay of curative intent treatment for ICC and is associated with improved survival in selected patients. We previously published data on a large, multi-institutional, multinational cohort of patients and noted that 5-year survival after surgical resection was 30% to 35%.

As with other cancers, prognostic information after surgical management of ICC is important to physicians and patients. Patients are interested in information about their overall survival based on factors specific to their individual case. Accurate prognostic systems for ICC are also important to physicians to help inform discussions around adjuvant therapy, as well as to stratify patients appropriately for clinical trials.

Prognostic systems for ICC are in a state of evolution. Initially staged the same as HCC, ICC only recently has had a separate, unique staging system introduced in the seventh edition of the American Joint Committee on Cancer (AJCC) Cancer Staging Manual. Although staging systems such as the AJCC system are helpful, such staging schemas are generally applicable to populations of patients rather than individual patients. Disease-specific nomograms have been proposed as a means to predict long-term survival in daily practice for the individual patient. Nomograms have been developed for a number of malignant neoplasms, including pancreatic adenocarcinoma, esophageal adenocarcinoma, and HCC. Although nomograms can provide important patient-level data on prognosis, they can be methodologically challenging to create.

Recently, Wang et al proposed a nomogram to predict survival after partial hepatectomy for ICC. The nomogram was based, however, on data from only one institution. In addition, the nomogram modeled the effect of several continuous variables (eg, age and tumor size) as linear coefficients. For example, tumor diameter was modeled as having a linear and constant effect on survival: as tumor size increased, survival was modeled as consistently getting worse. Methodologically, this may be problematic because data often should be transformed in nomogram models to predict more accurately the effect of these variables on survival. Previous studies have suggested that, rather than tumor size having a constant linear effect on outcome, there may be a threshold effect of tumor size on long-term outcomes, albeit at different tumor sizes.

The objective of the current study was to create a nomogram to predict long-term survival after surgical resection for ICC using individual patient-level factors. In addition, we more closely examined the effect of various potential prognostic factors, such as tumor size, to better assess the influence of such factors on prognosis. In creating and validating a proposed nomogram using our large, multi-institutional, international database, we sought to define a widely applicable and generalizable nomogram for use in patients with ICC.

Methods

From May 1, 1990, through December 31, 2011, a total of 514 patients with nonmetastatic ICC treated primarily with curative intent surgical resection were identified from 13 major hepatobiliary centers in the United States (The Johns Hopkins School of Medicine, Baltimore, Maryland; Duke Medical Center, Durham, North Carolina; University of Virginia Medical Center, Charlottesville; University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania; Massachusetts General Hospital, Boston; and Medical College of Wisconsin, Milwaukee), Europe (University Hospital Essen, Essen, Germany; Institute for Digestive Diseases and Liver Transplantation Fundeni, Bucharest, Romania; Curry Cabral Hospital, Lisbon, Portugal; Ospedale San Raffaele, Milan, Italy; Cliniques Universitaires Saint-Luc, Brussels, Belgium; and Hôpitaux Universitaires de Genève, Geneva, Switzerland), and Asia (Eastern Hepatobiliary Surgery Hospital, Shanghai, China). The study was approved by the institutional review boards of the respective institutions and informed consent was waived. Only patients with histologically confirmed ICC who received their initial treatment for ICC at a study center were included in the study.

Data Collection

The process of development of the study cohort has previously been described in detail. Briefly, patients were evaluated with a baseline history and physical examination, serum laboratory tests, and appropriate imaging studies (eg, computed tomography or magnetic resonance imaging of the abdomen and pelvis and radiography or computed tomography of the chest) at the discretion of the treating physician. Tumor characteristics were noted on imaging. Data on treatment-related variables, such as type of surgery, receipt of lymphadenectomy, and adjuvant therapy, were also obtained. Resection was classified as less than hemihepatectomy, hemihepatectomy, extended hepatectomy, or central hepatectomy. Margin, nodal status, vascular, and perineural infiltration were ascertained based on final pathologic assessment. The tumors were morphologically classified using the Liver Cancer Study Group of Japan criteria and staged using the TNM classification. After surgery, all patients were followed up routinely.

Date of last follow-up, vital status, and recurrence-related information were collected on all patients. We excluded patients with metastatic disease at the time of surgery or R2 margins on pathologic review. In addition, we excluded patients who died within 90 days of surgery to avoid including deaths due to postoperative complications within the overall survival statistics.

Statistical Analysis

Demographic, clinicopathologic, and treatment characteristics were described as summary statistics obtained using established methods and presented as percentages or median values. Time to death or censoring was estimated using the Kaplan-Meier method, and differences were compared using the log-rank test. All statistical tests were 2-tailed.

For purposes of constructing the nomogram, we aimed for a main model with fewer than 10 predictors. All clinically rel-
Relevant variables from the database of patients were considered. Tumor stage was considered as both the AJCC stage and individual variables included in the AJCC stage classification. Continuous predictors were transformed using cubic splines aiming to maximize the Wald χ² statistic. Within-hospital clustering of patients was accounted for in the model. Multiple potential interactions were tested. Factors were entered into the multivariate Cox proportional hazards regression model using backward stepwise selection with the Akaike information criterion (AIC), and coefficients of the predictors were calculated. Hazard ratios (HRs) and 95% CIs were estimated. Model performance was evaluated through discrimination indicated by the Harrell C-index, plotting the Kaplan-Meier curves of the quartiles of predictions, and further illustrated by drawing calibration plots using a bootstrapped sample. Model validation was performed using bootstrap resampling to quantify the overfitting of our modeling strategy and predict future performance of the model. Statistical analyses were performed using SAS, version 9.3 (SAS Institute Inc), and R, version 2.15.3, software packages (http://www.r-project.org/).

Results

Clinicopathologic and Treatment Characteristics
Of the 514 patients included in the study cohort, approximately half were male (273 [53.1%]) and white (314 [61.1%]). The median age of the patients was 59.2 years (Table 1). Patients were treated with an extended hepatectomy (202 [39.3%]), a hemihepatectomy (180 [35.0%]), or a minor liver resection (<3 segments) (132 [25.7%]) (Table 2). Most patients had a solitary tumor (384 [74.7%]). Macroscopic vascular invasion was noted in 56 patients (10.9%) and microscopic invasion in 68 (13.2%). Lymph node status was available for 252 patients; 90 patients (37.5%) had N1 disease. Most patients had early T1-T2-category tumors (451 [87.7%]). The median tumor size was 6.0 cm. Most patients had tumors 5 cm or larger (288 [56.0%]). Most patients had R0 surgical margins (452 [87.9%]), and 62 (12.1%) had R1 margins. Postoperatively, 122 (23.7%) patients received adjuvant therapy as part of their treatment, with 56 (45.9%) receiving a gemcitabine-based therapy.

The median survival for the cohort was 38.8 months (95% CI, 31.4-45.7 months). The 1-, 3-, and 5-year survival was 81.0%, 51.8%, and 39.8%, respectively.

Model Specification and Predictors of Survival
Initially, 15 clinically relevant candidate variables were selected from the database: age at diagnosis, sex, CA19-9 levels, and
carcinoembryonic antigen levels, hepatitis B virus status, hepatitis C virus status, presence of cirrhosis of the underlying liver, tumor size, nodal status, multifocality, histologic findings, vascular invasion, perineural invasion, biliary invasion, and direct invasion of contiguous organs (Table 3). Backward stepwise selection using the AIC in Cox proportional hazards regression modeling identified 6 variables that were the most associated with survival: age, tumor size, multiple lesions, nodal status, vascular invasion, and presence of cirrhosis of the underlying liver. Table 3 presents the HRs and 95% CIs for the multivariate Cox proportional hazards regression analysis for variables selected by the AIC. Age (HR, 1.33; 95% CI, 1.12-1.58), presence of multiple tumors (HR, 1.94; 95% CI, 1.50-2.52), tumor size (HR, 1.60; 95% CI, 1.18-2.17), underlying cirrhosis (HR, 1.24; 95% CI, 0.79-1.93), nodal status (HR, 1.78; 95% CI, 1.25-2.55), and macroscopic vascular invasion (HR, 2.36; 95% CI, 1.68-3.39) were independently associated with mortality (all P < .05), while the presence of cirrhosis tended to be associated with prognosis (HR, 1.51; 95% CI, 0.95-2.41) (P = .08).

Factors selected:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Univariate Model</th>
<th>Multivariate Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.33 (1.12-1.58)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Multiple tumors</td>
<td>1.94 (1.50-2.52)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Tumor size</td>
<td>1.60 (1.18-2.17)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Underlying cirrhosis</td>
<td>1.24 (0.79-1.93)</td>
<td>.19</td>
</tr>
<tr>
<td>Nodal status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N0</td>
<td>1.0 [Reference]</td>
<td>.01</td>
</tr>
<tr>
<td>N1</td>
<td>1.71 (1.21-2.40)</td>
<td>1.78 (1.25-2.55)</td>
</tr>
<tr>
<td>Nx</td>
<td>1.19 (0.92-1.44)</td>
<td>1.29 (0.97-1.73)</td>
</tr>
<tr>
<td>Vascular invasion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>1.0 [Reference]</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Microvascular</td>
<td>1.04 (0.82-1.46)</td>
<td>.94 (0.65-1.38)</td>
</tr>
<tr>
<td>Macrovascular</td>
<td>2.36 (1.68-3.39)</td>
<td>2.10 (1.46-3.01)</td>
</tr>
</tbody>
</table>

Factors not selected:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Univariate Model</th>
<th>Multivariate Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>1.11 (0.87-1.41)</td>
<td>.42</td>
</tr>
<tr>
<td>HCV infection</td>
<td>0.86 (0.43-1.74)</td>
<td>.67</td>
</tr>
<tr>
<td>HBV infection</td>
<td>1.23 (0.56-1.95)</td>
<td>.22</td>
</tr>
<tr>
<td>CEA, every 100-ng/mL increase</td>
<td>0.78 (0.51-1.20)</td>
<td>.26</td>
</tr>
<tr>
<td>CA19-9, every 100-U/mL increase</td>
<td>1.01 (1.00-1.01)</td>
<td>.06</td>
</tr>
<tr>
<td>Perineural invasion</td>
<td>1.32 (0.83-2.11)</td>
<td>.25</td>
</tr>
<tr>
<td>Biliary invasion</td>
<td>1.39 (0.93-2.07)</td>
<td>.11</td>
</tr>
<tr>
<td>Direct invasion of adjacent organs</td>
<td>1.81 (0.96-3.42)</td>
<td>.06</td>
</tr>
<tr>
<td>Moderate to well-differentiated tumor</td>
<td>0.71 (0.48-1.12)</td>
<td>.07</td>
</tr>
</tbody>
</table>

Abbreviations:

CEA, carcinoembryonic antigen; HBV, hepatitis B virus; HCV, hepatitis C virus.

Nomogram

An nomogram model to predict survival of patients with ICC undergoing surgical resection is shown in Figure 2. The nomogram was developed based on the 6 independent prognostic markers. In the nomogram model (Table 3), each factor was ascribed a weighted point total that implied a survival prognosis. For example, lymph node metastasis (N1) was associated with 11 points, whereas the presence of macroscopic vascular invasion was associated with 15 points.

Each patient with a higher score had a worse prognosis. Prognostic discrimination was performed by dividing the predicted survival probabilities into quartiles that were then used to plot Kaplan-Meier curves (Figure 3). Patients in the lowest quartile of predicted survival had a median recorded survival of 14.8 months, whereas those in the top quartile had a median recorded survival of 80.2 months (P = .01). The nomogram was able to stratify patients into 4 distinct incremental 5-year prognostic groups (quartile 1, 64.6%; quartile 2, 50.0%; quartile 3, 31.0%; and quartile 4, 15.4% [P = .01]).

Model Performance

Predictive accuracy (discrimination) of the final model was measured by calculating the Harrell C index, which was 0.692
Figure 4 presents the 40-sample bootstrapped calibration plot for the nomogram for prediction of 5-year survival. The calibration plots revealed good predictive accuracy of the nomogram (Figure 4). Bootstrap validation of the model with 300 iterations revealed minimal evidence of model overfit. The training data set C statistic was 0.699, and the testing data set C statistic was 0.706, which represented the bias-corrected estimate of model performance in the future. In comparison, the C statistics for ICC based on the sixth and seventh editions of the AJCC Cancer Staging Manual were 0.54 and 0.59, respectively.

Globally, ICC is the second most common primary liver malignant tumor. Frequently, ICC presents as advanced disease, and despite curative intent surgical resection, recurrence is common and less than 40% of patients survive longer than 5 years.\textsuperscript{1-3,11} Accurate information on prognostication is important for decision making and counseling of patients. Patients are interested in knowing the chance of cure and the odds of long-term survival based on data from their individual case. Widely used prognostications systems, such as the AJCC TNM classification, include a limited number of tumor-related variables and lack flexibility in terms of allowing physicians to tailor prognostication for specific patients. Nomograms are statistical tools that evaluate multiple factors and can predict the probability of long-term survival for individual patients. Prognostic nomograms have been proposed for a number of cancers.\textsuperscript{16-18} The current study is important because it defines a prognostic nomogram for long-term survival after curative intent surgery for ICC in a large, multinational, multi-institutional cohort of patients. We found that extremes of age, larger tumor diameter, macrovascular invasion, cirrhosis of the underlying liver, lymph node metastasis, and presence of multifocal disease adversely influenced survival based on the nomogram (Figure 2). More important, we also found that the association of tumor size and age on survival was not linear in nature and subsequently took this nonlinear relationship into account in the creation of the nomogram.

Nomograms are relatively easy to use, can evaluate a combination of predictors, and help clinicians better predict an individual patient’s outcome. Nomograms can incorporate factors that current staging systems do not consider and may show better individual discrimination than the AJCC staging system—the current standard for prognostication.\textsuperscript{10,16-18} Recently, Wang et al\textsuperscript{10} published the first nomogram for predicting long-term survival after hepatectomy for ICC and compared it with established staging systems. Although the nomogram proposed by Wang et al performed better than the seventh edition of the AJCC Cancer Staging Manual as well as the classification scheme proposed by de Jong et al\textsuperscript{3} and the Liver Cancer Study Group of Japan classification,\textsuperscript{4,14} the study was based on patients treated at a single institution in China. In addition, certain fac-
tors, such as age and tumor size, were modeled as having a "pure" linear effect on survival. Previous reports\(^5,19\) have noted, however, that estimating the effect of certain continuous variables using only a linear model may not accurately reflect the effect of the variable on outcome. In fact, transformation of continuous variables and evaluation of their association without outcome is a requisite step in multivariate modeling, especially for prediction nomograms.\(^5\) As such, we specifically sought to assess the effect of continuous predictors by transforming these data using cubic splines to maximize the Wald \(\chi^2\) statistic.

Interestingly, there are conflicting data on tumor size as a prognostic factor. Previous data from our group reported that tumor size did not have a notable effect on survival, and therefore tumor size was not incorporated into the TNM staging system for ICC.\(^4,20\) Subsequently, in examining recurrence rates and patterns of recurrence after surgery for ICC, we noted that tumor size seemed to affect risk of recurrence.\(^11\) The reason for this discrepancy is probably multifactorial. Certain factors may predict initial risk of recurrence, whereas others may be related to the subsequent clinical course and progression of the disease, leading to death.\(^21\) More important, data from the current study also suggested an alternative reason for the previous disparate results. Specifically, the prognostic importance of tumor size was found to have a nonlinear threshold effect on prognosis. The effect of tumor size on the risk of death was linear until an approximately 7-cm diameter, after which the risk of death associated with any further incremental increase in size plateaued (Figure 1A). A threshold effect of tumor size is noted for many other tumors and forms the basis of various cutoff values for tumors in the AJCC TNM classification.\(^4\) The failure of our group to identify an effect of tumor size on overall survival previously may be explained in part by the relative rarity of patients with ICC who presented with earlier-stage disease and small ICC tumors. Given that few patients had small tumors in the previously examined cohorts, we were probably less able to delineate the effect of tumor size on overall survival after resection of ICC.

In addition to tumor size, several other factors were selected based on multivariate analyses, including age at diagnosis, tumor number, vascular invasion, lymph node metastasis, and presence of underlying liver cirrhosis. All these factors had a strong effect on long-term survival and were therefore included in the final prediction nomogram model. Similar to tumor size, age was associated with a nonlinear effect on long-term survival (Figure 1B). Although worse survival among older patients undergoing curative intent surgery for ICC has been previously noted,\(^21\) the current study better defines the nonlinear relation of age to prognosis. Data in the current study also corroborated the effect of tumor number\(^20-24\) and lymph node metastasis\(^7\) on survival and quantitated the actual effect that these factors had on survival. Collectively, by aggregating predictive factors into a nomogram, we were able to stratify patients with good prognostic discrimination (Figure 3) and predictive accuracy (C-index, 0.692). In addition, the nomogram performed well on bootstrap validation with a training data set C statistic of 0.699 and a testing data set C statistic of 0.706, as well as calibration plots that showed good predictive accuracy of the nomogram (Figure 4).
The current study had several limitations. The data were derived from a large number of hepatobiliary centers from around the world. In addition, as with all retrospective studies of surgical procedures, the current cohort may have been subject to selection bias. Moreover, the patients were heterogeneous in demographic, clinical, and tumor-related characteristics. This heterogeneity can, however, be viewed as advantageous in that it allowed us to examine a “real-world” cohort, thereby facilitating generalizability of our findings. Although we performed rigorous validation of the nomogram using bootstrapped calibration and bias-corrected estimates, future studies will need to externally validate the proposed nomogram.

Conclusions

Extremes of age, incremental tumor diameter up to 7 cm, lymph node metastasis, vascular invasion, multifocality of lesions, and cirrhosis of the underlying liver predicted long-term survival and were incorporated into a prognostic nomogram. The nomogram was able to stratify patients into distinct prognostic groups and performed well on internal validation. Future studies should externally validate the proposed nomogram to establish more fully its value in the clinical prognostication of patients with ICC after surgical resection.

REFERENCES