Retrospective Cohort Study

INTERNATIONAL JOURNAL OF SURGERY

OPEN

Validation of prognostic models for predicting postsurgical outcomes in intrahepatic cholangiocarcinoma patients using a multicenter cohort

Dong Hwan Kim, MD, PhD[a], Sang Hyun Choi, MD, PhD[a,*], Sehee Kim, PhD[b], Woohyung Lee, MD, PhD[c], Hyung-Don Kim, MD, PhD[d], Hyungjin Rhee, MD, PhD[e], Eun-Suk Cho, MD[f], Suk-Keu Yeom, MD, PhD[g], Sumi Park, MD[h], Seung Soo Lee, MD, PhD[a], Mi-Suk Park, MD, PhD[e]

Background: Few studies to date have externally validated and directly compared conventional prognostic models that predict postsurgical outcomes in patients with intrahepatic cholangiocarcinoma (ICCA). We aimed to validate the performance of prognostic models predicting postsurgical outcomes in a multicenter cohort of patients with ICCA.

Materials and methods: Consecutive patients with ICCA who underwent curative-intent hepatic resection for ICCA at six tertiary referral institutions between January 2009 and December 2016 were retrospectively analyzed. The predictive abilities of the American Joint Committee on Cancer TNM 8th edition, Wang nomogram, Hyder nomogram, MEGNA score, and Tsilimigras prescore and postscore models were assessed based on preoperative and postoperative clinical, pathological, and imaging data. The ability of the six prognostic models to predict early recurrence (within 1-year) and 5-year overall survival (OS) was evaluated, including their discrimination and calibration capabilities.

Results: Among the 333 patients included (mean age [standard deviation], 62.4 [9.8] years; 206 [61.9%] men), 142 (42.6%) experienced early recurrence and 225 (67.6%) died during a median follow-up of 45.1 months. In predicting early recurrence, the Tsilimigras postscore model showed the highest area under the receiver operating characteristic curve (0.811), followed by the Wang nomogram (0.780). In predicting 5-year OS, the Wang nomogram had the highest concordance index (0.704), followed by the Tsilimigras postscore model (0.675). All six models showed good calibration, with strong agreement between predicted and observed outcomes. Risk stratification based on these models effectively differentiated OS at 1, 3, and 5 years post-surgery ($P \leq 0.005$).

Conclusion: The Tsilimigras postscore model and the Wang nomogram were optimal in predicting early recurrence and long-term survival, respectively, in ICCA patients who underwent curative-intent resection. These findings may help select the most appropriate prognostic model for predicting postsurgical outcomes in ICCA patients.

Keywords: cholangiocarcinoma, liver neoplasms, prognosis, recurrence, survival

Introduction

Intrahepatic cholangiocarcinoma (ICCA) is the second most common type of primary liver cancer, with its incidence and mortality rates rising globally[1–4]. The only potentially curative treatment is complete surgical resection, which offers a median overall survival (OS) of approximately 30 months[5]. Even after curative resection, however, patient prognosis remains poor, with 5-year OS rates of only 25%–40% and tumor recurrence rates as high as 50%–70%[3,5]. Notably, most recurrences occur early in the postoperative period, with patients experiencing early recurrence having significantly poorer prognoses than

[a]Department of Radiology and Research Institute of Radiology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea, [b]Department of Clinical Epidemiology and Biostatistics, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea, [c]Division of Hepatobiliary and Pancreatic Surgery, Department of Surgery, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea, [d]Department of Oncology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea, [e]Department of Radiology, Research Institute of Radiological Science, Severance Hospital, Yonsei University College of Medicine, Seoul, Republic of Korea, [f]Department of Radiology, Gangnam Severance Hospital, Yonsei University College of Medicine, Seoul, Republic of Korea, [g]Department of Radiology, Korea University Ansan Hospital, Korea University College of Medicine, Ansan, Republic of Korea and [h]Department of Radiology, National Health Insurance Service Ilsan Hospital, Goyang, Republic of Korea

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

Corresponding author. Address: Department of Radiology and Research Institute of Radiology, Asan Medical Center, University of Ulsan College of Medicine, 88, Olympic-Ro 43-Gil, Songpa-Gu, Seoul 05505, Republic of Korea. Tel.: +82 230101797; Fax: +82 24764719. E-mail: edwardchoi83@gmail.com (S. H. Choi).

International Journal of Surgery (2025) 111:7032–7040

Received 5 February 2025; Accepted 7 June 2025

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.lww.com/international-journal-of-surgery.

Published online 24 June 2025

http://dx.doi.org/10.1097/JS9.0000000000002810

patients with late recurrence[5–8]. Accurately predicting postoperative prognosis and identifying patients at high risk for early recurrence are therefore crucial for guiding individualized postoperative surveillance or considering alternative treatment strategies, such as neoadjuvant therapy or nonsurgical options.

The American Joint Committee on Cancer (AJCC) 8th edition TNM staging system is widely used in the pathological staging of ICCA. This system incorporates factors such as tumor size, multiplicity, lymphovascular invasion, lymph node metastasis, and distant metastasis, effectively stratifying the extent of disease in most patients. Survival outcomes in some patients, however, may not necessarily align with their AJCC TNM stages. Additionally, several studies have identified other prognostic factors, including surgical margin status, tumor differentiation, and patient age[5,9].

The limitations of the AJCC TNM staging system have led to the development of several other prognostic models, such as the Wang nomogram[10], the Hyder nomogram[11], the MEGNA score[12], and the Tsilimigras risk scoring system[7], to predict postoperative prognosis in patients with ICCA. Although these models incorporate additional variables and may offer greater accuracy than the conventional AJCC staging system, few specifically address the risk of clinically significant early recurrence, which may indicate the need for alternative treatment options[13]. Furthermore, some of these models are limited by relying on single-center-based cohorts or the lack of external validation[10,11]. Direct comparisons of their predictive performance are also constrained, as individual studies have evaluated patient cohorts with varying baseline characteristics. Therefore, the present study was designed to validate and directly compare the performance of these six prognostic models for predicting postsurgical outcomes in a multicenter cohort of patients with ICCA.

HIGHLIGHTS

- Six prognostic models for intrahepatic cholangiocarcinoma were externally validated using multicenter cohort data.
- Tsilimigras postscore showed the highest accuracy for early recurrence prediction.
- Wang nomogram was the most accurate for predicting 5-year overall survival.
- All models showed good calibration and effective postsurgical risk stratification.

Methods

This cohort study has been reported in line with the STROCSS guidelines[14], and was retrospectively registered in the Research Registry (UIN: researchregistry11256; accessible at https://www.researchregistry.com/browse-the-registry/#home/registrationdetails/683f02270197c0030d04a1b4/).

Study population

This multicenter, retrospective study was approved by the institutional review boards of six tertiary referral institutions in South Korea, which waived the requirement for patient-informed consent. The study included consecutive patients who underwent curative-intent hepatic resection for ICCA at one of the participating institutions between January 2009 and December 2016. Patients were included if they (a) were aged ≥18 years and (b) had undergone hepatic resection with a confirmed diagnosis of ICCA. Patients were excluded if they (a) had a history of other malignancies; (b) underwent hepatic resection with palliative intent; (c) had another type of ICCA, such as intraductal papillary neoplasm of the bile duct (IPNB); (d) had received preoperative neoadjuvant treatment for ICCA; (e) had not undergone preoperative contrast-enhanced magnetic resonance imaging (MRI) or computed tomography (CT) within 3 months prior to surgery; (f) had missing data on serum concentrations of carbohydrate antigen 19-9 (CA19-9) or carcinoembryonic antigen (CEA), which are required for analysis using the Wang nomogram[10]; (g) did not have available pathologic data on tumor differentiation or microvascular invasion, which are necessary for analysis using the MEGNA score[12] and Tsilimigras risk score systems[7]; or (h) had died or were lost to follow-up within 1 year after resection without evidence of recurrence (Fig. 1).

Clinical, pathological, and imaging data

Clinical and pathological data were retrospectively collected from the electronic medical records of each institution. Clinical data included demographic characteristics; concentrations of serum tumor markers (CA19-9 and CEA) and viral markers (hepatitis B surface antigen [HBsAg] and hepatitis C antibody); the extent of hepatectomy, and adjuvant treatment status. Major hepatectomy was defined as the resection of three or more Couinaud segments, whereas minor hepatectomy was defined as the resection of two or fewer segments or a wedge resection[15]. Pathological data included tumor size, number, and location; morphological type (mass-forming, intraductal growth, or periductal infiltrating); tumor differentiation; perineural invasion; microvascular or macrovascular invasion; visceral peritoneal invasion; extrahepatic organ invasion; resection margin status; lymph node status; presence of cirrhosis; and biliary intraepithelial neoplasia.

Given MRI's higher sensitivity than CT in detecting tumor multiplicity[16], preoperative contrast-enhanced MRI served as the primary imaging modality, with CT used only if MRI results were unavailable. The anonymized MRI and CT scans were independently reviewed by three board-certified abdominal radiologists, each with over 13 years of experience in hepatic imaging, who were blinded to clinical, laboratory, pathological, and prognostic information. Imaging findings were analyzed specifically for inclusion in the Tsilimigras preoperative model, which incorporates tumor size, tumor number, and the presence of underlying liver cirrhosis (Supplemental Digital Content Table S1, available at: http://links.lww.com/JS9/E454)[7]. Tumor size and number were determined by calculating the mean of measurements provided by the three reviewers, whereas liver cirrhosis was determined by the consensus of at least two reviewers.

Assessment of prognostic models

TNM stage was determined from pathologic data according to the eighth edition of the AJCC system (Supplemental Digital

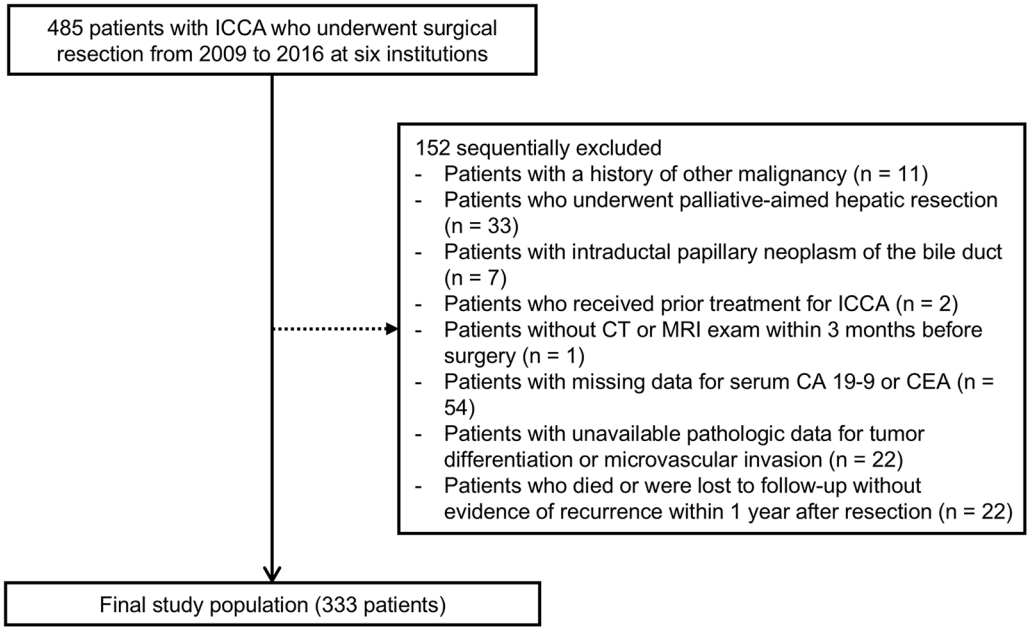

Figure 1. Flow diagram of the study population. ICCA, intrahepatic cholangiocarcinoma; MRI, magnetic resonance imaging; CT, computed tomography; CA 19-9, carbohydrate antigen 19-9; CEA, carcinoembryonic antigen.

Content Table S2, available at: http://links.lww.com/JS9/E454)[17]. Positive lymph node metastasis was classified as pathologic N1 stage, regardless of the number of harvested regional lymph nodes, whereas negative lymph node metastasis was classified as pathologic N0 stage if at least six regional lymph nodes had been completely sampled. Patients negative for lymph node metastasis but with incomplete lymph node sampling, as well as those who did not undergo lymphadenectomy, were classified as having pathologic Nx stage. N stage for these patients was determined using composite reference standards, with the development of suspicious regional lymph nodes on follow-up CT, MRI, or positron emission tomography (PET) within 3 months after surgery classified as N1; absence of suspicious regional lymph nodes on follow-up imaging for at least 1 year after surgery classified as N0; and all others classified as Nx[16].

Other prognostic models, including the Wang nomogram[10], the Hyder nomogram[11], the MEGNA score[12], and the Tsilimigras risk scoring system[7], were retrospectively evaluated. The preoperative and postoperative variables required by each prognostic model are summarized in Supplemental Digital Content Tables S3, available at: http://links.lww.com/JS9/E454. Except for the Tsilimigras risk scoring system[7], which was designed to predict very early recurrence (within 6 months), all other models were designed for long-term survival following resection. Because two Tsilimigras risk scoring systems were developed based on preoperative and postoperative characteristics[7], these two models called the Tsilimigras prescore and postscore models, respectively, were validated separately.

Postoperative outcomes

Postoperative patient outcomes were obtained from their electronic medical records and/or cancer registry data at each participating institution. After surgical resection, patients underwent routine follow-up evaluations, including contrast-enhanced CT or MRI scans and measurement of the serologic tumor markers (CA19-9 and CEA) every 3–6 months. Patients showing increases in tumor marker concentrations and inconclusive imaging findings were evaluated by additional chest CT and/or PET-CT scans. Recurrence was defined as either suspicious or positive findings on follow-up imaging or histologically confirmed disease. The current status of each patient (alive, deceased, or lost to follow-up) and the number of days survived after surgery were assessed. Early recurrence was defined as recurrence within 1 year after resection[8,18–23]. OS was defined as the time interval from the date of surgery to the date of death from any cause. Patients who did not experience the event of interest were censored at the time of the last available follow-up.

Statistical analysis

The primary endpoint was early recurrence (within 1 year), whereas the secondary endpoint was the 5-year OS rate following surgery. The performance of the six prognostic models was validated by determining their discrimination and calibration abilities. Discrimination was defined as the ability to differentiate patients who did and did not experience an event (i.e. early recurrence or 5-year mortality), with early recurrence assessed by determining the area under the receiver operating characteristic (ROC) curve (AUC) and 5-year OS assessed by determining Harrell's C-index. Specifically, time-dependent ROC analysis was performed to determine 1, 3, and 5-year OS rates. Models were compared using a significance test based on the bootstrap method, with Bonferroni correction applied to adjust for multiple pairwise comparisons (i.e. $\alpha = 0.05/15$). Calibration refers to the degree of agreement between model-derived and observed probabilities, with the results presented as calibration plots. Calibration performance was further evaluated using the Hosmer–Lemeshow test or the Grønnesby and Borgan test.

Additionally, the Brier score of each model was calculated to summarize prediction error and provide an overall measure of model performance. Subgroup analyses of predictive performance were conducted based on the patient enrollment period (2009–2012 versus 2013–2016) and receipt of adjuvant therapy. OS was estimated using the Kaplan–Meier method and compared across risk groups using the log-rank test. The cutoff values for risk groups in the AJCC TNM stage, MEGNA score, and Tsilimigras risk scoring models were those defined in their original publications. By contrast, risk groups for the Wang and Hyder nomograms were determined based on the first and third quartiles of their risk scores, as their original publications did not explicitly define risk group categories[10,11].

Statistical analyses were performed using SPSS version 29.0 (IBM Corp.) or R version 4.3.2 (R Foundation for Statistical Computing), with $P < 0.05$ indicating statistical significance. For multiple comparisons, Bonferroni corrected P-values were calculated.

Results

Patient characteristics and follow-up

Of the 452 eligible patients, 119 were excluded for the following reasons: 11 had a history of other malignancies, 33 underwent palliative resection, 7 had IPNB, 2 had received preoperative treatment for ICCA, 1 lacked preoperative MRI or CT results, 54 had missing serum CA 19-9 or CEA data, 22 had unavailable pathological data on tumor differentiation or microvascular invasion, and 22 either died or were lost to follow-up within 1 year without recurrence (Fig. 1). The final study cohort included 333 patients, consisting of 206 (61.9%) men and 127 (38.1%) women, of mean age 62.4 ± 9.8 years; their clinical, pathological, and imaging characteristics are listed in Table 1. Of these patients, 206 (61.9%) underwent major hepatectomy, and 290 (87.1%) had negative surgical resection margins. Among 224 patients with pathologic Nx stage, 11 were reclassified as N1 and 183 as N0 based on the composite reference standards. The final N stage distribution was 90 patients classified as N1, 213 as N0, and 30 as Nx. Postoperatively, 150 patients (45.0%) received adjuvant chemotherapy, of whom 92 (61.3%) received gemcitabine-based regimens, and 8 (5.3%) received additional radiation therapy. A comparison of baseline characteristics between the study cohort and patients excluded due to missing serum CA 19-9 or CEA data is presented in Supplemental Digital Content Table S4, available at: http://links.lww.com/JS9/E454. No significant differences were observed between the groups, except for HBsAg positivity.

During a median follow-up period of 45.1 months (interquartile range: 18.7–115.0 months), 142 (42.6%) patients experienced early recurrence, and 225 (67.6%) died, of whom 199 (88.4%) were attributed to tumor progression. Of the 142 patients who experienced early recurrence, 45 (31.7%) experienced intrahepatic recurrence alone, 63 (44.4%) had extrahepatic recurrence alone, and 34 (23.9%) had both. The liver was the most common site of early recurrence ($n = 79$), followed by the lymph nodes ($n = 48$), peritoneum ($n = 35$), lungs ($n = 24$), bones ($n = 4$), abdominal wall ($n = 3$), adrenal gland ($n = 1$), and ovary ($n = 1$). The cumulative OS rates at 1, 3, and 5 years were 83.8%, 55.6%, and 42.9%, respectively.

Table 1

Baseline characteristics of the study populations

	Total ($n = 333$)
Clinical characteristics	
Age, mean (SD), year	62.4 (9.8)
Sex	
Male	206 (61.9)
Female	127 (38.1)
HBsAg (+)	56 (16.8)
Anti-HCV (+)	10 (3.0)
Serum CA 19-9, U/mL	37.7 (11.9–274.0)
Serum CEA, μg/L	2.7 (1.7–5.1)
Resection procedure	
Major hepatectomy	206 (61.9)
Minor hepatectomy	127 (38.1)
Adjuvant chemotherapy[a]	150 (45.0)
Pathologic characteristics	
Tumor size, mean (SD), cm	5.2 (2.8)
Tumor multiplicity	37 (11.1)
Macroscopic feature	
MF type, purely	307 (92.2)
MF combined with the PI or IG type	26 (7.8)
Differentiation	
Well or moderate	254 (76.3)
Poor	79 (23.7)
Microvascular invasion	176 (52.9)
Macrovascular invasion	28 (8.4)
Perineural invasion	116 (34.8)
Cirrhosis	29 (8.7)
Biliary intraepithelial neoplasia	30 (9.0)
Positive resection margin	43 (12.9)
Imaging characteristics	
Preoperative imaging modality	
Magnetic resonance imaging	312 (93.7)
Computed tomography	21 (6.3)
Tumor size, mean (SD), cm	4.9 (2.5)
Tumor number, mean (SD)	1.3 (1.1)
Radiographic evidence of cirrhosis	44 (13.2)

Unless specified otherwise, values are number (%) of patients or median (interquartile range).
SD, standard deviation; HBsAg, hepatitis B surface antigen; HCV, hepatitis C virus; CA, carbohydrate antigen; CEA, carcinoembryonic antigen; MF, mass-forming; PI, periductal infiltrative; IG, intraductal growing; AJCC, American Joint Committee on Cancer.
[a]Eight patients received additional radiation therapy.

According to the AJCC TNM staging system, 107 patients (32.1%) were classified as stage I, 89 (26.7%) as stage II, 130 (39.1%) as stage III, and 7 (2.1%) as stage IV (Table 2). The Wang nomogram classified 83 (24.9%), 166 (49.9%), and 84 (25.2%) into low (score ≤27.19), intermediate (score 27.20–111.88), and high (score >111.88) risk groups, respectively; whereas the Hyder nomogram classified 83 (24.9%), 169 (50.8%), and 81 (24.3%) patients into low (score ≤8.75), intermediate (score 8.76–14.18), and high (score >14.18) risk groups, respectively. The MEGNA score stratified patients into four risk groups: low (score 0; $n = 73$, 21.9%), intermediate (score 1; $n = 138$, 41.4%), high (score 2; $n = 91$, 27.3%), and very high risk (score ≥3; $n = 31$, 9.3%). The Tsilimigras prescore classified 182 (54.7%), 103 (30.9%), and 48 (14.4%) patients into low (score ≤1.77), intermediate (score 1.78–2.69), and high (score >2.69) risk groups, respectively, whereas the Tsilimigras postscore classified 184 (55.3%), 90 (27.0%), and 59 (17.7%) patients into low (score ≤2.02), intermediate (score 2.03–2.80), and high (score >2.80) risk groups, respectively.

Table 2

Summary of risk score according to prognostic models

Prognostic model	Value
TNM stage, AJCC 8th edition	
I	107 (32.1)
II	89 (26.7)
III	130 (39.1)
IV	7 (2.1)
Wang nomogram, median (interquartile range)[a]	54.28 (27.19–111.88)
Low risk (≤27.19)	83 (24.9)
Intermediate risk (27.20–111.88)	166 (49.9)
High risk (>111.88)	84 (25.2)
Hyder nomogram, median (interquartile range)[a]	11.78 (8.76–14.18)
Low risk (≤8.75)	83 (24.9)
Intermediate risk (8.76–14.18)	169 (50.8)
High risk (>14.18)	81 (24.3)
MEGNA score	
Low risk (0)	73 (21.9)
Intermediate risk (1)	138 (41.4)
High risk (2)	91 (27.3)
Very high risk (≥3)	31 (9.3)
Tsilimigras prescore, median (interquartile range)	1.68 (1.29–2.32)
Low risk (≤1.77)	182 (54.7)
Intermediate risk (1.78–2.69)	103 (30.9)
High risk (>2.69)	48 (14.4)
Tsilimigras postscore, median (interquartile range)	1.92 (1.38–2.51)
Low risk (≤2.02)	184 (55.3)
Intermediate risk (2.03–2.80)	90 (27.0)
High risk (>2.81)	59 (17.7)

Otherwise specified, data are number of patients with the percentage in parentheses.
AJCC, American Joint Committee on Cancer; MEGNA, multifocality, extrahepatic extension, grade, node positivity, and age older than 60 years.
[a]Risk groups were classified based on the first and third quartiles of the risk scores.

Predictive performances for early recurrence

Table 3 provides a pairwise comparison of the discriminative abilities of the six prognostic models to predict early recurrence. The Tsilimigras postscore achieved the highest AUC (0.811, 95% confidence interval [CI]: 0.764–0.859), somewhat higher than the Wang nomogram (0.780, 95% CI: 0.729–0.830, $P = 0.150$), but significantly higher than the other models ($P \leq 0.001$). By contrast, the MEGNA score had the lowest AUC (0.651, 95% CI: 0.593–0.709), significantly lower than both the Tsilimigras postscore and Wang nomogram ($P < 0.001$). Calibration plots showed overall good agreement between the predicted and observed risks of early recurrence across all six models (Supplemental Digital Content Figure S1 and Tables S5 and S6, available at: http://links.lww.com/JS9/E454). The prediction error, measured by the Brier score, was lowest for the Tsilimigras postscore (0.177) and highest for the MEGNA score (0.225) (Table 3). Subgroup analyses based on patient enrollment period and receipt of adjuvant therapy yielded consistent results (Supplemental Digital Content Table S7 available at: http://links.lww.com/JS9/E454), with the Tsilimigras postscore maintaining the highest AUC across all subgroups.

Predictive performances for 5-year OS

Among the six prognostic models, the Wang nomogram had the highest C-index (0.704, 95% CI: 0.670–0.738), followed by the Tsilimigras postscore (0.675, 95% CI: 0.642–0.709) (Table 3). The C-index of the Wang nomogram was significantly higher than that of the MEGNA score (0.622, 95% CI: 0.580–0.663) and Tsilimigras prescore (0.579, 95% CI: 0.539–0.619) ($P < 0.001$). Calibration plots indicated overall good correlation between the predicted and observed risks of 5-year mortality across all six models (Supplemental Digital Content Figure S1 and Tables S5 and S6, available at: http://links.lww.com/JS9/E454). Although the Brier scores of the six models were generally similar, the Wang nomogram had the lowest Brier score, 0.203 (Table 3). The time-dependent AUC with 95% CI and integrated Brier scores at 1-, 3-, and 5-years are presented in Supplemental Digital Content Table S8, available at: http://links.lww.com/JS9/E454. Although the time-dependent AUC remained stable over time, the integrated Brier scores gradually increased across all models, indicating an increase in prediction error over time. OS at 1, 3, and 5 years after surgery was significantly different according to the risk groups determined by each prognostic model ($P \leq 0.005$), although the risk stratification by the Tsilimigras prescore appeared less distinct (Fig. 2 and Supplemental Digital Content Table S9, available at: http://links.lww.com/JS9/E454). Subgroup analyses based on patient enrollment period and receipt of adjuvant therapy showed consistent results (Supplemental Digital Content Table S7, available at: http://links.lww.com/JS9/E454), with the Wang nomogram consistently demonstrating the highest C-index across all subgroups. Risk stratification by each model within subgroups based on adjuvant therapy use is shown in Supplemental Digital Content Figure S2, available at: http://links.lww.com/JS9/E454.

Discussion

This study compared the performance of six prognostic models for predicting postsurgical outcomes, including early recurrence and 5-year mortality in ICCA patients using data from a multicenter cohort. As none of the enrolled patients had missing values, direct comparisons between the prognostic models were feasible. In intra-individual comparisons, the Tsilimigras postscore model showed the highest ability to predict early recurrence, whereas the Wang nomogram showed the highest ability to predict 5-year OS. All six models exhibited good calibration and effective risk stratification capabilities. Our study provides a comprehensive evaluation of prognostic models in a large, multicenter Asian cohort and includes subgroup analyses to assess performance across clinically relevant factors.

The Tsilimigras postscore model was found to be the optimal prognostic model for predicting early recurrence (AUC, 0.811). This model significantly outperformed the other models, including the AJCC TNM stage, Hyder nomogram, MEGNA score, and Tsilimigras prescore. The Tsilimigras postscore model, which was developed using a cohort of 880 ICCA patients who underwent resection at several international institutions[7], included several independent predictive factors, such as patient age, race, tumor size, tumor number, microvascular invasion, resection margin status, and lymph node metastasis. As our study cohort consisted entirely of nonwhite patients, race was not a contributing factor in our analysis. Consequently, the model's performance may have been underestimated, further supporting its generalizability across racial and geographic populations. Notably, unlike the other models, the Tsilimigras postscore includes surgical margin status, a well-established prognostic factor in ICCA[9,24]. Because positive surgical margins are independently associated with adverse prognostic

Table 3

Comparison of predictive performance of prognostic models

		P-value[a] for comparison with					
		AJCC TNM	Wang nomogram	Hyder nomogram	MEGNA score	Tsilimigras prescore	Brier score
Early recurrence	**AUC (95% CI)**						
AJCC TNM	0.721 (0.669–0.773)						0.205
Wang nomogram	0.780 (0.729–0.830)	0.014					0.196
Hyder nomogram	0.732 (0.678–0.786)	0.697	0.053				0.208
MEGNA score	0.651 (0.593–0.709)	0.024	<0.001	0.007			0.225
Tsilimigras prescore	0.692 (0.634–0.749)	0.433	0.011	0.258	0.332		0.220
Tsilimigras postscore	0.811 (0.764–0.859)	<0.001	0.150	0.001	<0.001	<0.001	0.177
5-year mortality	**Harrell's C-index (95% CI)**						
AJCC TNM	0.658 (0.624–0.693)						0.211
Wang nomogram	0.704 (0.670–0.738)	0.006					0.203
Hyder nomogram	0.664 (0.622–0.705)	0.780	0.016				0.219
MEGNA score	0.622 (0.580–0.663)	0.065	<0.001	0.033			0.230
Tsilimigras prescore	0.579 (0.539–0.619)	<0.001	<0.001	<0.001	0.139		0.238
Tsilimigras postscore	0.675 (0.642–0.709)	0.334	0.073	0.557	0.027	<0.001	0.211

[a]*P*-values <0.003 (0.05/15) are considered statistically significant after applying the Bonferroni correction for multiple comparisons.

AUC, area under the receiver operating characteristic curve; CI, confidence interval; AJCC, American Joint Committee on Cancer; MEGNA, multifocality, extrahepatic extension, grade, node positivity, and age older than 60 years.

outcomes[25], their inclusion likely played a significant role in enhancing the model's predictive performance in the postoperative setting. Additionally, this model incorporated microvascular invasion status, a recently recognized adverse prognostic indicator for ICCA[26,27]. In contrast to other models focused on long-term survival, the Tsilimigras risk scoring system was designed to predict very early recurrence (within 6 months), with the present study confirming that the Tsilimigras postscore model can be useful in predicting early recurrence. Considering that adjuvant chemotherapy with capecitabine alone has demonstrated a survival benefit in biliary tract cancer, including ICCA[28], the Tsilimigras postscore may be particularly valuable for identifying patients who are more likely to benefit from adjuvant chemotherapy.

The Wang nomogram was found to be optimal in predicting 5-year OS (C-index, 0.704) in patients with ICCA. This model significantly outperformed the MEGNA score and Tsilimigras prescore models. The Wang nomogram model was derived from a single-center cohort of 367 Asian patients who underwent resection for ICCA. Of the six prognostic models for ICCA, the Wang nomogram was the only model that included serum tumor markers, which have been independently associated with OS in ICCA patients[5,29]. The prognostic value of these markers is further supported by subsequent studies showing that they can improve ICCA staging and survival prediction[30–33], which may have contributed to the superior performance of the Wang nomogram. Additionally, its inclusion of direct extrahepatic invasion, a factor also included in the AJCC TNM staging system, may reflect aggressive tumor biology. The superior prognostic performance of the Wang nomogram is in partial agreement with findings showing that it performed better than the AJCC TNM staging system and the Hyder nomogram[34]. However, because the Wang nomogram classifies lymph node metastasis into two categories, presence or absence, patients with unknown lymph node status (Nx) were classified as node-negative (N0). Therefore, this model may be more appropriate for patients undergoing formal portal lymphadenectomy.

By contrast, the Tsilimigras prescore exhibited suboptimal performance for predicting both early recurrence (AUC, 0.692) and 5-year OS (C-index, 0.579). Although current evidence does not support the routine use of neoadjuvant chemotherapy over upfront resection in resectable ICCAs[35,36], neoadjuvant therapy may be considered in select cases, such as patients with advanced disease or those at high risk of early progression[13]. It may also facilitate tumor downsizing and downstaging in initially unresectable ICCAs[37]. Of the six prognostic models evaluated in the present study, the Tsilimigras prescore was the only model based solely on preoperative variables, making it uniquely suited to inform decisions regarding neoadjuvant therapy. However, the predictive performance of this model was suboptimal. Therefore, it may be unreliable to base clinical decisions on the Tsilimigras prescore alone. This highlights the inherent trade-off between model accessibility and predictive accuracy. In resource-limited settings or during preoperative planning – where postoperative pathological data may not be available – reliance on preoperative models remains necessary despite their limitations. Therefore, there is a clear need to develop and validate more accurate preoperative prognostic models that balance feasibility with performance. Ultimately, model selection should be tailored to the clinical context, taking into account the timing of decision-making, data availability, and institutional resources.

Recent advances, including radiomics and deep learning-based models, have shown significant potential for accurately predicting early recurrence[18,20–23]. Integrating these approaches into prognostic modeling represents a promising direction for future research. Additionally, the extraction and analysis of data from liquid biopsy, including circulating tumor DNA, enable the detection of subtle tumor characteristics that may not be captured by conventional clinical or pathological parameters. Incorporating these data sources into existing prognostic models may improve predictive accuracy, facilitate dynamic risk assessment over time, and enable more personalized treatment

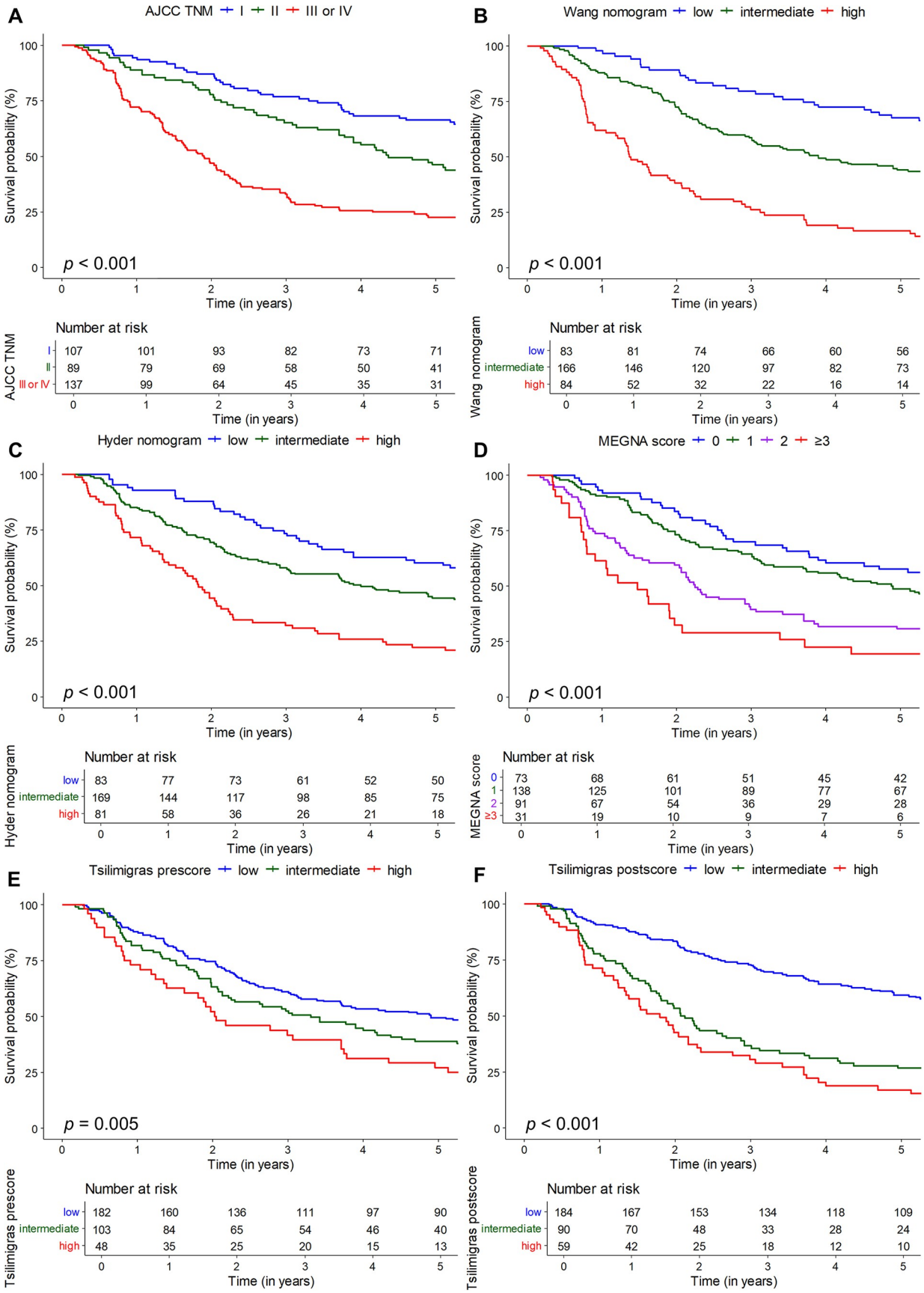

Figure 2. Kaplan–Meier analyses of 5-year overall survival across risk groups stratified by AJCC TNM stage (A), the Wang nomogram (B), the Hyder nomogram (C), MEGNA score (D), Tsilimigras prescore (E), and Tsilimigras postscore (F).

approaches, including timely interventions such as the selection of adjuvant therapy for high-risk patients and monitoring of treatment response[38,39].

The present study had several limitations. First, the dataset was collected retrospectively, which may introduce biases inherent to this approach. Second, to enable intra-individual comparisons between models, only patients with complete datasets were included, leading to the exclusion of a substantial number of patients. This may have introduced selection bias. However, because no significant differences were observed in baseline characteristics or model performance before and after excluding patients with missing data (Supplemental Digital Content Table S10, available at: http://links.lww.com/JS9/E454), this limitation is unlikely to have substantially affected the results. Third, although current guidelines recommend a 6-month course of adjuvant chemotherapy following curative resection of ICCA[3,13,28], only 45.0% of patients in the present cohort received adjuvant chemotherapy, potentially affecting tumor recurrence and patient survival. Notably, these patients were enrolled between 2009 and 2016, a period during which robust evidence supporting the use of adjuvant chemotherapy had not yet been established through randomized clinical trials. However, the percentage of patients who received adjuvant chemotherapy in the present study (45.3%) was higher than the percentages reported in recent studies for ICCA patients (18.9%–36.3%)[6,7,21,23]. Fourth, although some models have been validated in previous studies[34,40], this study offers unique contributions, including the use of multicenter data, detailed subgroup analyses, and an evaluation of racial differences affecting the performance of the Tsilimigras models. Lastly, while this study addresses a gap in the existing literature by providing a direct comparison of prognostic models, it does not propose a new model, and thus the overall level of innovation may be considered limited.

In conclusion, the Tsilimigras postscore and Wang nomogram demonstrated superior performance in predicting early recurrence and long-term survival, respectively, in ICCA patients who underwent curative-intent resection, outperforming other prognostic models. These findings from direct comparisons of models, combined with a comprehensive understanding of the patient and tumor biologic factors incorporated into each model, may improve the precision and personalization of postoperative care by guiding the selection of the most appropriate prognostic model for predicting postsurgical outcomes in patients with ICCA. Future prospective studies involving international patient cohorts are warranted to validate our findings and to support the development of novel prognostic models that incorporate advanced techniques such as artificial intelligence.

Ethical approval

This study was approved by the institutional review board of all six participating institutions.

Consent

The need for informed consent was waived by the institutional review board of all participating institutions.

Sources of funding

There was no funding support.

Author contributions

All authors had access to the study data and had reviewed and approved the final manuscript. Study concept/design: D.H. K. and S.H.C.; data acquisition: S.H.C., H.R., E.-S.C., S.K.Y., S.P., S.L., and M.S.P; data analysis and interpretation: D.K., S. C., S.K.Y., S.S.L., and M.S.P.; drafting of the manuscript: D.H. K.; critical revision of the manuscript for important intellectual content: all authors; statistical analysis: D.H.K., S.K., and S.H. C.; study supervision: S.H.C.

Conflicts of interest disclosure

The authors declared that they have no conflicts of interest.

Research registration unique identifying number (UIN)

researchregistry11256.

Guarantor

Sang Hyun Choi is the guarantor of this study.

Provenance and peer review

Not commissioned, externally peer-reviewed.

Data availability statement

The datasets analyzed during the current study are not publicly available due to the datasets containing information that could compromise the privacy of research participants, but are available from the corresponding author on reasonable request.

References

[1] Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. CA Cancer J Clin 2024;74:12–49.
[2] Bertuccio P, Malvezzi M, Carioli G, *et al*. Global trends in mortality from intrahepatic and extrahepatic cholangiocarcinoma. J Hepatol 2019;71:104–14.
[3] European Association for the Study of the Liver. European Association for the Study of the Liver. EASL-ILCA clinical practice guidelines on the management of intrahepatic cholangiocarcinoma. J Hepatol 2023;79: 181–208.
[4] Konyn P, Ahmed A, Kim D. The current trends in the health burden of primary liver cancer across the globe. Clin Mol Hepatol 2023;29: 358–62.
[5] Mavros MN, Economopoulos KP, Alexiou VG, Pawlik TM. Treatment and prognosis for patients with intrahepatic cholangiocarcinoma: systematic review and meta-analysis. JAMA Surg 2014;149:565–74.
[6] Zhang XF, Beal EW, Bagante F, *et al*. Early versus late recurrence of intrahepatic cholangiocarcinoma after resection with curative intent. Br J Surg 2018;105:848–56.
[7] Tsilimigras DI, Sahara K, Wu L, *et al*. Very early recurrence after liver resection for intrahepatic cholangiocarcinoma: considering alternative treatment approaches. JAMA Surg 2020;155:823–31.

[8] Wang C, Pang S, Si-Ma H, *et al*. Specific risk factors contributing to early and late recurrences of intrahepatic cholangiocarcinoma after curative resection. World J Surg Oncol 2019;17:2.
[9] Sahara K, Tsilimigras DI, Mehta R, *et al*. A novel online prognostic tool to predict long-term survival after liver resection for intrahepatic cholangiocarcinoma: the "metro-ticket" paradigm. J Surg Oncol 2019;120:223–30.
[10] Wang Y, Li J, Xia Y, *et al*. Prognostic nomogram for intrahepatic cholangiocarcinoma after partial hepatectomy. J Clin Oncol 2013;31:1188–95.
[11] Hyder O, Marques H, Pulitano C, *et al*. A nomogram to predict long-term survival after resection for intrahepatic cholangiocarcinoma: an Eastern and Western experience. JAMA Surg 2014;149:432–38.
[12] Raoof M, Dumitra S, Ituarte PHG, *et al*. Development and validation of a prognostic score for intrahepatic cholangiocarcinoma. JAMA Surg 2017;152:e170117.
[13] Moris D, Palta M, Kim C, Allen PJ, Morse MA, Lidsky ME. Advances in the treatment of intrahepatic cholangiocarcinoma: an overview of the current and future therapeutic landscape for clinicians. CA Cancer J Clin 2023;73:198–222.
[14] Agha R, Mathew G, Rashid R, *et al*. Revised Strengthening the Reporting of Cohort, Cross-Sectional and Case-Control Studies in Surgery (STROCSS) Guideline: an update for the age of artificial intelligence. Prem J Sci 2025;10:100081.
[15] Strasberg SM. Nomenclature of hepatic anatomy and resections: a review of the Brisbane 2000 system. J Hepatobiliary Pancreat Surg 2005;12:351–55.
[16] Kim YY, Yeom SK, Shin H, *et al*. Clinical staging of mass-forming intrahepatic cholangiocarcinoma: computed tomography versus magnetic resonance imaging. Hepatol Commun 2021;5:2009–18.
[17] Lee AJ, Chun YS. Intrahepatic cholangiocarcinoma: the AJCC/UICC 8th edition updates. Chin Clin Oncol 2018;7:52.
[18] Liang W, Xu L, Yang P, *et al*. Novel nomogram for preoperative prediction of early recurrence in intrahepatic cholangiocarcinoma. Front Oncol 2018;8:360.
[19] Li Q, Zhang J, Chen C, *et al*. A nomogram model to predict early recurrence of patients with intrahepatic cholangiocarcinoma for adjuvant chemotherapy guidance: a multi-institutional analysis. Front Oncol 2022;12:896764.
[20] Wakiya T, Ishido K, Kimura N, *et al*. CT-based deep learning enables early postoperative recurrence prediction for intrahepatic cholangiocarcinoma. Sci Rep 2022;12:8428.
[21] Alaimo L, Lima HA, Moazzam Z, *et al*. Development and validation of a machine-learning model to predict early recurrence of intrahepatic cholangiocarcinoma. Ann Surg Oncol 2023;30:5406–15.
[22] Song Y, Zhou G, Zhou Y, *et al*. Artificial intelligence CT radiomics to predict early recurrence of intrahepatic cholangiocarcinoma: a multicenter study. Hepatol Int 2023;17:1016–27.
[23] Bo Z, Chen B, Yang Y, *et al*. Machine learning radiomics to predict the early recurrence of intrahepatic cholangiocarcinoma after curative resection: a multicentre cohort study. Eur J Nucl Med Mol Imaging 2023;50:2501–13.
[24] Farges O, Fuks D, Boleslawski E, *et al*. Influence of surgical margins on outcome in patients with intrahepatic cholangiocarcinoma: a multicenter study by the AFC-IHCC-2009 study group. Ann Surg 2011;254:824–29.
[25] Li MX, Bi XY, Li ZY, *et al*. Impaction of surgical margin status on the survival outcome after surgical resection of intrahepatic cholangiocarcinoma: a systematic review and meta-analysis. J Surg Res 2016;203:163–73.
[26] Ali SM, Clark CJ, Mounajjed T, *et al*. Model to predict survival after surgical resection of intrahepatic cholangiocarcinoma: the Mayo Clinic experience. HPB (Oxford) 2015;17:244–50.
[27] Hu LS, Weiss M, Popescu I, *et al*. Impact of microvascular invasion on clinical outcomes after curative-intent resection for intrahepatic cholangiocarcinoma. J Surg Oncol 2019;119:21–29.
[28] Primrose JN, Fox RP, Palmer DH, *et al*. Capecitabine compared with observation in resected biliary tract cancer (BILCAP): a randomised, controlled, multicentre, phase 3 study. Lancet Oncol 2019;20:663–73.
[29] Bergquist JR, Ivanics T, Storlie CB, *et al*. Implications of CA19-9 elevation for survival, staging, and treatment sequencing in intrahepatic cholangiocarcinoma: a national cohort analysis. J Surg Oncol 2016;114:475–82.
[30] Wang Z, Shi Y, Xiong G, Han M, Chen X. The prognostic impact of preoperative CA19-9 on resectable cholangiocarcinoma: a comprehensive systematic review and meta-analysis. Discov Oncol 2024;15:773.
[31] Sasaki K, Margonis GA, Andreatos N, *et al*. Serum tumor markers enhance the predictive power of the AJCC and LCSGJ staging systems in resectable intrahepatic cholangiocarcinoma. HPB (Oxford) 2018;20:956–65.
[32] He C, Zhang Y, Song Y, *et al*. Preoperative CEA levels are supplementary to CA19-9 levels in predicting prognosis in patients with resectable intrahepatic cholangiocarcinoma. J Cancer 2018;9:3117–28.
[33] Loosen SH, Roderburg C, Kauertz KL, *et al*. CEA but not CA19-9 is an independent prognostic factor in patients undergoing resection of cholangiocarcinoma. Sci Rep 2017;7:16975.
[34] Doussot A, Groot-Koerkamp B, Wiggers JK, *et al*. Outcomes after resection of intrahepatic cholangiocarcinoma: external validation and comparison of prognostic models. J Am Coll Surg 2015;221:452–61.
[35] Wehrle CJ, Chang J, Woo K, *et al*. Neoadjuvant therapy reduces node positivity but does not confer survival benefit versus up-front resection for resectable intrahepatic cholangiocarcinoma: a propensity-matched analysis. J Surg Oncol 2024;130:453–61.
[36] Van Treeck BJ, Olave MC, Watkins RD, *et al*. Neoadjuvant therapy leads to objective response in intrahepatic cholangiocarcinoma. HPB (Oxford) 2024;26:938–48.
[37] Maithel SK, Keilson JM, Cao HST, *et al*. NEO-GAP: a single-arm, phase II feasibility trial of neoadjuvant gemcitabine, cisplatin, and Nab-Paclitaxel for resectable, high-risk intrahepatic cholangiocarcinoma. Ann Surg Oncol 2023;30:6558–66.
[38] Yoo C, Jeong H, Jeong JH, *et al*. Circulating tumor DNA status and dynamics predict recurrence in patients with resected extrahepatic cholangiocarcinoma. J Hepatol 2025;82:861–70.
[39] Zanuso V, Tesini G, Valenzi E, Rimassa L. New systemic treatment options for advanced cholangiocarcinoma. J Liver Cancer 2024;24:155–70.
[40] Hahn F, Muller L, Mahringer-Kunz A, *et al*. Risk prediction in intrahepatic cholangiocarcinoma: direct comparison of the MEGNA score and the 8th edition of the UICC/AJCC cancer staging system. PLoS One 2020;15:e0228501.