

## Retrospective Study

**Estimating survival benefit of adjuvant therapy based on a Bayesian network prediction model in curatively resected advanced gallbladder adenocarcinoma**

Zhi-Min Geng, Zhi-Qiang Cai, Zhen Zhang, Zhao-Hui Tang, Feng Xue, Chen Chen, Dong Zhang, Qi Li, Rui Zhang, Wen-Zhi Li, Lin Wang, Shu-Bin Si

**ORCID number:** Zhi-Min Geng (0000-0003-2645-9808); Zhi-Qiang Cai (0000-0002-7380-8110); Zhen Zhang (0000-0003-4150-1410); Zhao-Hui Tang (0000-0002-2903-8008); Feng Xue (0000-0003-2941-3733); Chen Chen (0000-0002-7916-7940); Dong Zhang (0000-0001-8254-7568); Qi Li (0000-0001-7139-8819); Rui Zhang (0000-0002-1727-5524); Wen-Zhi Li (0000-0002-6990-3430); Lin Wang (0000-0002-9594-3989); Shu Bin Si (0000-0003-2297-4423).

**Author contributions:** Geng ZM, Cai ZQ, Tang ZH, and Si SB designed the research; Xue F accessed the SEER database and acquired the data; Geng ZM, Cai ZQ, Zhang Z, Tang ZH, Chen C, Zhang D, Li Q, Zhang R, Li WZ, Wang L, and Si SB analyzed and interpreted the data; Geng ZM, Zhang Z, and Cai ZQ drafted the manuscript; Tang ZH and Si SB revised the manuscript critically; Geng ZM and Cai ZQ contributed equally to this work.

**Supported by** the National Natural Science Foundation of China, No. 81572420 and No. 71871181; the Key Research and Development Program of Shaanxi Province, No. 2017ZDXM-SF-055; and the Multi-center Clinical Research Project of School of Medicine, Shanghai Jiaotong University, No. DLY201807.

**Institutional review board statement:** This study was reviewed and approved by the

**Zhi-Min Geng, Feng Xue, Chen Chen, Dong Zhang, Qi Li, Rui Zhang, Wen-Zhi Li, Lin Wang,** Department of Hepatobiliary Surgery, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an 710061, Shaanxi Province, China

**Zhi-Qiang Cai, Zhen Zhang, Shu-Bin Si,** Department of Industrial Engineering, School of Mechanical Engineering, Northwestern Polytechnical University, Xi'an 710072, Shaanxi Province, China

**Zhao-Hui Tang,** Department of General Surgery, Shanghai Xin Hua Hospital Affiliated to School of Medicine, Shanghai Jiaotong University, Shanghai 200092, China

**Corresponding author:** Shu-Bin Si, PhD, Dean, Professor, Department of Industrial Engineering, School of Mechanical Engineering, Northwestern Polytechnical University, 127 West Youyi Road, Xi'an 710072, Shaanxi Province, China. [sisb@nwpu.edu.cn](mailto:sisb@nwpu.edu.cn)  
**Telephone:** +86-13991363388

**Abstract****BACKGROUND**

The factors affecting the prognosis and role of adjuvant therapy in advanced gallbladder carcinoma (GBC) after curative resection remain unclear.

**AIM**

To provide a survival prediction model to patients with GBC as well as to identify the role of adjuvant therapy.

**METHODS**

Patients with curatively resected advanced gallbladder adenocarcinoma (T3 and T4) were selected from the Surveillance, Epidemiology, and End Results database between 2004 and 2015. A survival prediction model based on Bayesian network (BN) was constructed using the tree-augmented naïve Bayes algorithm, and composite importance measures were applied to rank the influence of factors on survival. The dataset was divided into a training dataset to establish the BN model and a testing dataset to test the model randomly at a ratio of 7:3. The confusion matrix and receiver operating characteristic curve were used to evaluate the model accuracy.

**RESULTS**

Institution Review Board of The First Affiliated Hospital of Xi'an Jiaotong University.

**Informed consent statement:** As this study is based on a publicly available database without identifying patient information, informed consent was not needed.

**Conflict-of-interest statement:** All authors declare no conflict of interest related to this article.

**Data sharing statement:** No additional data are available.

**Open-Access:** This article is an open-access article which was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

**Manuscript source:** Unsolicited manuscript

**Received:** July 4, 2019

**Peer-review started:** July 5, 2019

**First decision:** August 2, 2019

**Revised:** August 30, 2019

**Accepted:** September 9, 2019

**Article in press:** August 2, 2019

**Published online:** October 7, 2019

**P-Reviewer:** Uhlmann D, Wada R, Zhang XB

**S-Editor:** Wang J

**L-Editor:** Wang TQ

**E-Editor:** Ma YJ



A total of 818 patients met the inclusion criteria. The median survival time was 9.0 mo. The accuracy of BN model was 69.67%, and the area under the curve value for the testing dataset was 77.72%. Adjuvant radiation, adjuvant chemotherapy (CTx), T stage, scope of regional lymph node surgery, and radiation sequence were ranked as the top five prognostic factors. A survival prediction table was established based on T stage, N stage, adjuvant radiotherapy (XRT), and CTx. The distribution of the survival time (>9.0 mo) was affected by different treatments with the order of adjuvant chemoradiotherapy (cXRT) > adjuvant radiation > adjuvant chemotherapy > surgery alone. For patients with node-positive disease, the larger benefit predicted by the model is adjuvant chemoradiotherapy. The survival analysis showed that there was a significant difference among the different adjuvant therapy groups (log rank, surgery alone *vs* CTx,  $P < 0.001$ ; surgery alone *vs* XRT,  $P = 0.014$ ; surgery alone *vs* cXRT,  $P < 0.001$ ).

## CONCLUSION

The BN-based survival prediction model can be used as a decision-making support tool for advanced GBC patients. Adjuvant chemoradiotherapy is expected to improve the survival significantly for patients with node-positive disease.

**Key words:** Gallbladder carcinoma; Bayesian network; Surgery; Adjuvant therapy; Prediction model

©The Author(s) 2019. Published by Baishideng Publishing Group Inc. All rights reserved.

**Core tip:** A Bayesian network model was constructed to predict the survival time for patients with advanced gallbladder carcinoma (GBC) after curative resection from the Surveillance, Epidemiology, and End Results database, with a model accuracy of 69.67%, and the area under the curve for the testing dataset was 77.72%. Adjuvant radiation, chemotherapy, and T stage were ranked as the top three prognostic factors by importance measures. The prediction model supported the role of adjuvant therapy for advanced GBC patients after curative resection. Adjuvant chemoradiotherapy is expected to improve the survival more significantly for patients with node-positive disease.

**Citation:** Geng ZM, Cai ZQ, Zhang Z, Tang ZH, Xue F, Chen C, Zhang D, Li Q, Zhang R, Li WZ, Wang L, Si SB. Estimating survival benefit of adjuvant therapy based on a Bayesian network prediction model in curatively resected advanced gallbladder adenocarcinoma. *World J Gastroenterol* 2019; 25(37): 5655-5666

**URL:** <https://www.wjgnet.com/1007-9327/full/v25/i37/5655.htm>

**DOI:** <https://dx.doi.org/10.3748/wjg.v25.i37.5655>

## INTRODUCTION

Gallbladder carcinoma (GBC) is the sixth most common gastrointestinal malignancy and the most common biliary tract cancer worldwide, with a dismal 5-year survival rate of 5% to 10% in most cases<sup>[1-3]</sup>. Radical resection is the cornerstone of curative therapy for GBC. The factors such as the involvement of the liver<sup>[4]</sup>, lymphatic metastases<sup>[5]</sup>, and jaundice<sup>[6]</sup>, which influence the prognosis of patients with advanced GBC after radical resection, are still in dispute. The role of adjuvant therapy for GBC, however, is not well-known at this time<sup>[7-9]</sup>. Therefore, the identification of postoperative prognostic factors for patients with advanced GBC and the establishment of an accurate survival prediction model are of great significance for the selection of individualized treatments to increase the survival time.

Prediction models have been developed to assist doctors in estimating probabilities and potentially provide decision-making support<sup>[10]</sup>. Data mining methods have recently been introduced for use in the survival prediction of patients with GBC. Wang *et al*<sup>[11]</sup> built a nomogram survival model to predict which GBC patients may benefit from adjuvant chemoradiotherapy (cXRT). Additionally, Wang *et al*<sup>[12]</sup> proposed a multivariate Cox proportional hazards model to predict the survival

benefit of adjuvant radiotherapy (XRT) for GBC patients. Ethun *et al*<sup>[13]</sup> designed a novel, pathology-based preoperative risk score to predict locoregional residual and distant disease, and the survival for incidental GBC. However, these methods are unable to represent the variables under uncertainty and also ignore the cause-and-effect relationships between prognostic factors.

Bayesian network (BN) is a directed acyclic graph for probabilistic reasoning<sup>[14]</sup>. Combined with a machine learning algorithm, BNs have great advantages in exploring the unknown probability of variables from the known probability knowledge, which has been applied widely in the field of medicine<sup>[15,16]</sup>. We have previously applied a BN model and importance measures to identify the significant factors of survival after surgery for patients with GBC<sup>[17]</sup>. In the present study, we applied BN to build a model to predict the survival time for patients with advanced GBC following curative resection from the Surveillance, Epidemiology, and End Results (SEER) database. The objective of the study was to provide a survival prediction model to patients with GBC as well as to identify the role of adjuvant therapy.

## MATERIALS AND METHODS

### Data sources and study population

Patients who suffered GBC between 2004 and 2015 were identified from the SEER database, which included 12722 patients and 133 variables. First, only patients who were diagnosed based on the American Joint Committee on Cancer (AJCC) sixth and seventh editions from 2004 to 2015 were selected. We removed patients who did not have the ICD-0-3 codes 8140, 8141, 8143, or 8147, which designate adenocarcinoma<sup>[18]</sup>. Then, we chose GBC patients whose T stages were T3 and T4 and whose M stage was M0. Third, we removed patients who did not have Surg Prim Site codes 40 and 60, which indicate radical resection for GBC patients. Meanwhile, we selected patients whose behavior code was 3, which meant that the tumor was malignant. Finally, we removed patients with NA that did not incorporate the variables positive Ln and tumor size. After selection, we identified 818 patients who demonstrated advanced gallbladder adenocarcinoma and had undergone radical resection. The patient selection flowchart is shown in [Figure 1](#).

Standard patient data included 18 attribute variables, which were sex, age, grade, positive Ln, number Ln, Surg Prim Site, Scope Reg LN Sur, Surg Oth Reg, SEER historic stage, tumor size, T stage, N stage, AJCC stage, radiation, radiation sequence, chemotherapy, vital status, and survival months. The variables are shown in [Table 1](#).

### Prognostic model based on Bayesian network

Data from the original dataset including 818 patients and 18 variables were collected to establish a BN model. As the BN can only deal with standard and discrete data formats, standardization of the dataset has to be completed prior to modeling ([Table 1](#)). The variable of age was divided into three intervals of 19 to 64 years, 65 to 75 years, and 76 to 97 years based on the equal frequency. Positive Ln was divided into three intervals of 0, 1 to 3, and >3. Number Ln was divided into four intervals of 0, 1 to 3, 4 to 6, and >6. Tumor size was divided into four intervals of 0 to 10 mm, 11 to 30 mm, 31 to 50 mm, and >50 mm based on medical definitions. Survival time was divided into two intervals according to median survival time.

After establishment of the dataset of patients with advanced GBC and transformation of all continuous prognostic factors into discrete variables, a prognostic model was established using the tree-augmented naïve (TAN) Bayes algorithm. In order to evaluate the model performance more accurately, the strategy of stratification sampling was adopted to split the dataset to a training dataset and testing dataset. Seventy percent (574) of patients formed the training dataset to establish the Bayesian model and the remaining 30% (244) patients were considered as the testing dataset to test the model. With the dataset prepared, survival months were considered as the target variable to predict and the rest of variables were deemed to be the prognostic factors. The established BN model is shown in [Figure 2](#).

### Confusion matrix and receiver operating characteristic curve

For the confusion matrix, the reliability of a model is defined as the values along the major diagonal of the total instance. Meanwhile, partial reliability is calculated by positive prediction value (PPV) = true positive (TP)/[TP + false positive (FP)], true positive rate (TPR) = TP/[TP + false negative (FN)], and true negative rate (TNR) = true negative(TN)/[FP + true negative (TN)]. Model accuracy is defined by the following equation: (1) Accuracy = (TP + TN)/(TP + FP + TN + FN).

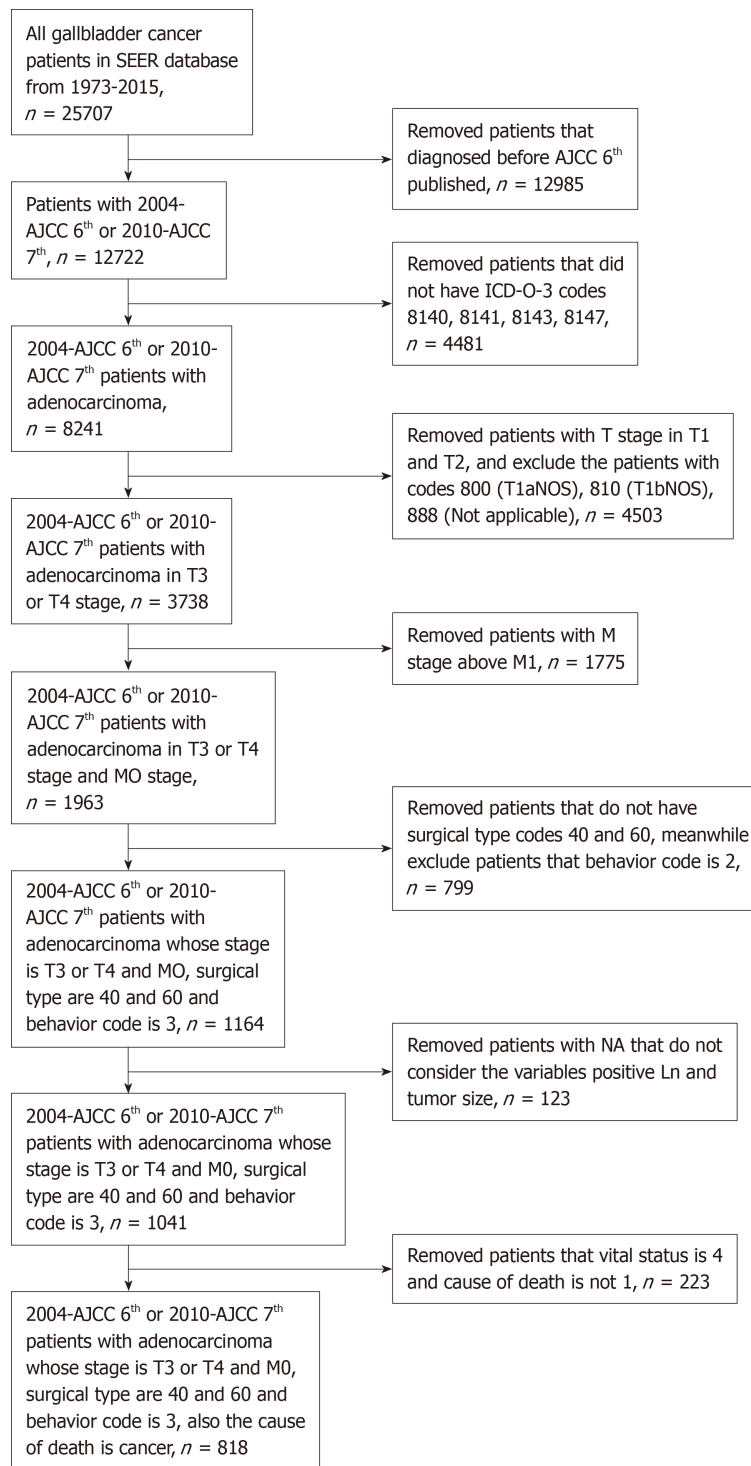


Figure 1 Patient selection flowchart.

When the number of negative and positive cases is imbalance in dataset, the accuracy may not be the appropriate criteria. Considering this condition, the receiver operating characteristic curve and the area under the curve (AUC), one of useful evaluating criteria, were calculated to measure the overall performance of the classification model further.

**Importance measure**

The Birnbaum importance measure is one of the most useful multistate importance measures available and could be used to calculate the importance of different attribute factors. The multistate Birnbaum importance equation can be expressed as:

(2) *formula*

where S means the survival time. Variables are described as (V1, V2,

**Table 1 Demographic and clinical characteristics of patients with advanced gallbladder adenocarcinoma and the calculation and rank of importance for prognostic factors in survival time**

Variable	Value	Frequency, <i>n</i> (%)	Priori probability (%)	Posterior probability (%)	Importance	Rank
Sex	1 (male)	261 (31.91)	33.26	48.69	0.0274	16
	2 (female)	557 (68.09)	66.72	54.05		
Age	1 (19-64)	256 (31.30)	30.84	39.55	0.1087	10
	2 (65-75)	258 (31.54)	31.71	42.31		
	3 (76-97)	304 (37.16)	37.46	56.74		
Grade	1 (well)	70 (8.56)	8.36	43.75	0.1383	8
	2 (moderately)	363 (44.38)	45.30	37.69		
	3 (poorly)	376 (45.97)	45.82	56.27		
	4 (undifferentiated)	9 (1.10)	0.52	66.67		
Positive Ln	0 (0)	172 (21.03)	65.26	47.93	0.0737	13
	1 (1-3)	229 (28.00)	30.47	43.19		
	2 (> 3)	27 (3.30)	4.27	56.86		
	NA	390 (47.68)				
Number Ln	0 (0)	389 (47.56)	46.52	56.93	0.1440	6
	1 (1-3)	266 (32.52)	34.49	40.19		
	2 (4-6)	76 (9.29)	8.01	34.78		
	3 (> 6)	87 (10.64)	10.98	31.75		
Surg Prim Site	40	627 (76.65)	77.00	49.32	0.1068	11
	60	191 (23.35)	23.00	38.64		
Scope Reg Ln Sur	0	387 (47.31)	46.69	57.09	0.1824	4
	1	5 (0.61)	0.52	33.33		
	3	7 (0.86)	1.05	16.67		
	4	264 (32.27)	33.80	40.72		
	5	153 (18.70)	17.94	33.98		
	6	1 (0.12)				
	7	1 (0.12)				
Surg Oth Reg	0	659 (80.56)	79.97	48.80	0.1205	9
	1	17 (2.08)	2.44	42.86		
	2	121 (14.79)	14.81	40.00		
	4	15 (1.83)	2.09	33.33		
	5	6 (0.73)	0.70	25.00		
Seer historic stage	1	197 (24.08)	23.87	48.91	0.0292	15
	2	514 (62.84)	62.72	45.56		
	4	107 (13.08)	13.41	49.35		
Tumor size	0 (0-10 mm)	20 (2.44)	3.49	39.69	0.0797	12
	1 (11-30 mm)	228 (27.87)	40.64	42.98		
	2 (31-50 mm)	181 (22.13)	31.19	44.46		
	3(> 50 mm)	132 (16.14)	24.68	57.31		
	NA	257 (31.42)				
T stage	300 (T3)	768 (93.89)	92.86	45.40	0.2045	3
	400 (T4)	50 (6.11)	7.14	65.85		
N stage	0 (N0)	534 (65.28)	64.98	47.99	0.0321	14
	100 (N1)	284 (34.72)	35.02	44.78		
AJCC stage	520 (stage 3A)	515 (62.96)	62.72	47.50	0.1423	7
	530 (stage 3B)	253 (30.93)	30.14	41.04		
	720 (stage 4A)	50 (6.11)	7.14	68.85		
Radiation	0 (no)	592 (72.37)	71.78	56.07	0.3261	1
	1 (yes)	226 (27.63)	28.22	23.46		
Radiation sequence	0 (no)	592 (72.37)	71.78	56.07	0.1805	5
	2 (radiation before surgery)	4 (0.49)	0.35	50.00		
	3 (radiation after surgery)	222 (27.14)	27.87	23.12		
Chemo-therapy	0 (no)	473 (57.82)	57.14	58.84	0.2795	2

	1(yes)	345 (42.18)	42.86	30.89
Vital status	1 (Survival)	197 (24.08)		
	4 (Dead)	621 (75.92)		
Survival months	≤ 9	383 (46.82)	46.86	
	> 9	435 (53.18)	53.14	

NA: Not available; Positive Ln: The exact number of regional lymph nodes examined by the pathologist that were found to contain metastases; Number Ln: The total number of regional lymph nodes that were removed and examined by the pathologist; Surg Prim Site: Surgery of primary site. 40: Total surgical removal of primary site; 60: Radical surgery; Scope Reg LN Sur: Scope of regional lymph node surgery; 0: No regional lymph nodes removed or aspirated; 1: Biopsy or aspiration of regional lymph node; 3: Number of regional lymph nodes removed unknown; not stated, 4: 1 to 3 regional lymph nodes removed, 5: 4 or more regional lymph nodes removed; 7: Sentinel node biopsy and code 3, 4, or 5 at different times; Surg Oth Reg: Surgical procedure of other site; 0: None, diagnosed at autopsy; 1: Nonprimary surgical procedure performed; 2: Nonprimary surgical procedure to other regional sites; 4: Nonprimary surgical procedure to distant site; 5: Combination of codes 2, 3, or 4; SEER historic stage: 1 Localized, 2 Regional, 3 Distant. AJCC: American Joint Committee on Cancer.

$V_3, \dots, V_i, \dots, V_n$ ) and each variable has states  $(0, 1, \dots, j, \dots, W_j)$ .  $P(S = 0)$  describes the prior probability of survival time  $\leq$  median survival time, while  $P(S = 0 | V_i = j)$  represents posterior probability.

### Statistical analysis

Statistical analyses were performed using the Statistical Package for the Social Sciences version 22.0 for Windows (IBM Corp., Armonk, NY, United States) and R software version 3.6.1 (<http://www.r-project.org/>). BayesiaLab (Bayesian Ltd. Co., France) was used to construct the BN model. Datasets were prepared using Microsoft Excel 2013 (Microsoft Corp., Redmond, WA, United States). All continuous variables were transformed to discrete variables for BN analysis and are expressed as frequency and percent. Categorical variables are presented as frequency and percent. Survival curve was estimated using the Kaplan–Meier method and the results were compared by the log-rank test.

## RESULTS

### General characteristics of the study population

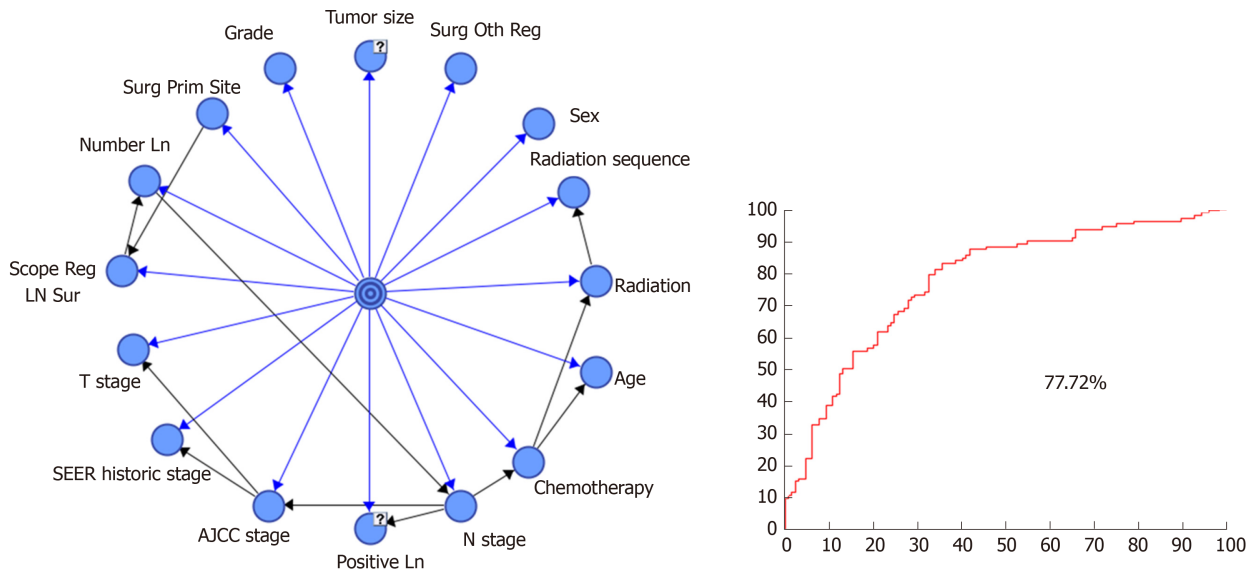
A total of 818 patients who underwent radical resection for advanced GBC met the inclusion criteria. The demographic and other characteristics of these patients are presented in [Table 1](#). The proportions of patients with T3 and T4 disease were 93.89% and 6.11%, respectively. Meanwhile, the proportions of patients with N0 and N1 disease were 65.28% and 34.72%, respectively. With regard to radiation, 226 (27.63%) patients received it, while 592 (72.37%) did not or had unclear data. Out of the patients who received radiation, only 4 (0.49%) received it before surgery, while 222 (27.14%) received it after surgery. Regarding chemotherapy, 345 (42.18%) underwent it, while 473 (57.82%) did not or had unclear data.

### Assessment of model efficacy

When the follow-up was cut off in December 2015, 621 patients were dead and 197 patients were still alive. The median survival time was 9.0 mo (interquartile, 4–19 mo). A total of 574 patients in the training dataset were used to establish the BN model, and the testing dataset including 244 patients was used to evaluate the model performance. The results of confusion matrix are listed in [Table 2](#). There were 114 patients whose survival time was less than 9 mo and 130 patients who had a survival time longer than 9 mo. A total of 84 patients were correctly classified as having survival time  $\leq$  9 mo and 86 patients were classified as having survival time  $>$  9 mo, based on the probability threshold of 0.5. Therefore, the TPR was 73.68%, while the precision or PPV was 65.62%. Thus, we inferred that the model accuracy was 69.67% based on Equation (1). The AUC value for the testing dataset was 77.72%, which reflects the accuracy of the model ([Figure 2](#)).

### Prognostic factors ranked by importance

After we established the BN model using the training dataset, we listed the priori probability of each variable and calculated the posterior probability according to the equation of  $P(S = 0 | V_i = j)$ . The priori probability for survival time was [ $P(S = 0) = 0.4686$ ,  $P(S = 1) = 0.5314$ ]. Then, the importance measure of each variable was calculated using Equation (2); the ranking results are shown in [Table 1](#). The results indicated that adjuvant radiation was the most important prognosis factor influencing survival time after radical resection for advanced GBC patients. Meanwhile, variables



**Figure 2** Bayesian network model for advanced gallbladder adenocarcinoma patients after radical resection and the receiver operating characteristic curve for survival time  $\leq 9$  mo of the Bayesian network model. “?” means that the variable has not available values, and we use Structural Expectation-maximization algorithm to impute the not available values in the dataset. Blue lines represent the relationship between the attribute variable and the target variable. Black lines represent the relationship between the attribute variables. BN: Bayesian network; SEER: Surveillance, Epidemiology, and End Results; AJCC: American Joint Committee on Cancer.

of chemotherapy and T stage played crucial roles in the prognosis of GBC patients, which ranked in the second and third places. Beyond that, the Scope Reg Ln Sur and radiation sequence were the fourth and fifth attribute factors for prognosis.

### Survival prediction model

We combined the BN model and importance measures to identify that radiation and chemotherapy were important prognostic factors. Meanwhile, T stage and N stage are always used to determine the severity of the patient's illness. As a result, we selected radiation, chemotherapy, T stage, and N stage as the observation variables to obtain a prediction table. Using the established BN model, we could predict the survival probability for a survival time  $> 9$  mo for patients with different states who underwent different adjuvant therapies. The survival prediction model is shown in Table 3.

For patients with node-negative disease, the model estimated the survival benefit from the addition of XRT and cXRT, regardless of T3 or T4 stage. For example, for a patient with T3N0 disease, his/her probability of a survival time  $> 9$  mo with surgery alone, adjuvant chemotherapy (CTx), XRT, and cXRT was 41.35%, 58.29%, 75.42%, and 76.62%, respectively. For patients with node-positive disease, the model predicted survival benefit from CTx, XRT and xCRT. Meanwhile, patients acquired more benefit from xCRT than CTx and XRT. For example, for a patient with T4N1 disease, his/her probability of a survival time  $> 9$  mo with surgery alone, CTx, XRT, and cXRT was 14.85%, 37.03%, 43.14%, and 57.97%, respectively.

### Survival analysis

The 1-, 3-, and 5-year overall survival (OS) rates for advanced GBC patients after radical resection were 47.5%, 17.8%, and 10.9%, respectively. The median OS for the advanced GBC patients with surgery alone, CTx, XRT, and cXRT was 6.0, 13.0, 15.0, and 21.0 mo, respectively. There was a significant difference among the different adjuvant therapy groups (log rank, surgery alone *vs* CTx,  $P < 0.001$ ; surgery alone *vs* XRT,  $P = 0.014$ ; surgery alone *vs* cXRT,  $P < 0.001$ ) (Figure 3).

## DISCUSSION

The use of an effective risk and survival prediction model in any cancer after surgical resection is critical to patients and physicians for making decisions regarding adjuvant therapy modalities and the frequency of follow-up<sup>[19]</sup>. However, the accurate prediction of cancer outcome remains one of the most interesting and challenging tasks for physicians. As a result, machine learning methods have become a popular

$$MBM_{i,j}^s = \frac{1}{w_j - 1} \sum_{j=1}^m |P(S=0 | V_i = j) - P(S=0)|$$

tool for medical researchers<sup>[20,21]</sup>.

In this study, we integrated BN with importance measures to identify the key factors underlying the prognosis of patients with advanced GBC after curative resection under uncertainty. The BN model was applied to predict patient survival time using the data gathered from the SEER database. In our model, the general model accuracy was 69.67%. Furthermore, the AUC for the BN model in the testing dataset was 77.72%. Therefore, we could obtain a higher TPR with a given FPR, which indicates a higher prediction accuracy with a lower risk.

On the basis of this model, by combining importance measures and their calculation method, we could obtain the ranking of the prognostic factors that influence the survival of patients with advanced GBC after radical resection. The ranking of prognostic factors is listed as follows: Radiation, chemotherapy, T stage, Scope Reg LN Sur, radiation sequence, Number Ln, AJCC stage, grade, Surg Oth Reg, age, Surg Prim Site, tumor size, positive Ln, N stage, SEER historic stage, and sex, in which radiation, chemotherapy, and T stage ranked among the top three with an importance value of 0.3261, 0.2795, and 0.2045, respectively. A total of 226 (27.63%) and 345 (42.18%) patients received radiation and chemotherapy, respectively. The results showed that adjuvant radiation and chemotherapy were the most important prognostic factors that influenced the survival of patients with advanced GBC after radical resection.

Locoregional recurrence following radical resection is a problem facing successful curing of the disease in patients with biliary tract cancers; for this reason, regional treatment in the form of radiotherapy is often considered as an adjuvant therapy after surgery<sup>[22]</sup>. However, due to its uncommon nature and the limited randomized controlled trials, the role of adjuvant therapy in advanced GBC patients after curative resection is controversial<sup>[23-25]</sup>.

A nomogram survival model proposed by Wang *et al.*<sup>[11]</sup> predicted that GBC patients with at least T2 or N1 disease would gain a survival benefit from adjuvant CRT, with a C-index of 0.67. Another multivariate Cox proportional hazards model constructed by Wang *et al.*<sup>[12]</sup> predicted that adjuvant RT provided a survival benefit in node-positive or  $\geq$  T2 disease, with a C-index of 0.71. An analysis from the National Cancer Data Base by Hoehn *et al.*<sup>[26]</sup> reported that adjuvant chemotherapy and radiation therapy improved survival for patients with node-positive GBC. In addition, a phase II prospective trial of adjuvant capecitabine and gemcitabine followed by radiotherapy and concurrent capecitabine in extrahepatic cholangiocarcinoma and GBC (SWOG: S0809) showed high levels of local control<sup>[27]</sup>.

In this study, the BN prediction model confirmed the role of adjuvant therapy in advanced GBC patients after curative resection. For patients with node-negative disease, the model estimated the survival benefit from the addition of XRT and cXRT, regardless of T3 or T4 stage. For patients with node-positive disease, the model predicted survival benefit from CTx, XRT, and xCRT. Meanwhile, patients acquired more benefit from xCRT than CTx and XRT. In addition, the predicted result was validated by the survival analysis.

To our knowledge, no similar research on adjuvant therapy in curative resected GBC based on BN was reported. BN is a directed acyclic graph for probabilistic reasoning and is different from the traditional statistical analysis. However, the current study has several limitations. There are two main limitations affect recommended analyses using the SEER radiotherapy and chemotherapy data: (1) The completeness of the variables; and (2) The biases associated with unmeasured reasons for receiving or not receiving RT/chemotherapy. In addition, we were unable to consider some of the details from the SEER database, including radiation dosage, field(s) of radiation, and margin status as well as details of chemotherapy. Moreover, although this study was performed using the SEER database, the study population was only 818 patients after screening. Large-volume, prospective, randomized, controlled clinical trials are therefore needed to validate the prediction model in the future.

In conclusion, a BN model was constructed to predict the survival time for patients with advanced GBC after curative resection from the SEER database, with a model accuracy of 69.67%, and the AUC value for the testing dataset was 77.72%. Importance measures were used to sort these prognostic factors, in which radiation, chemotherapy, and T stage were ranked among the top three. The prediction model supported the role of adjuvant therapy for advanced GBC patients after curative resection. For patients with node-negative disease, the model estimated the survival



**Table 2** Confusion matrix, reliability, and accuracy of Bayesian network model

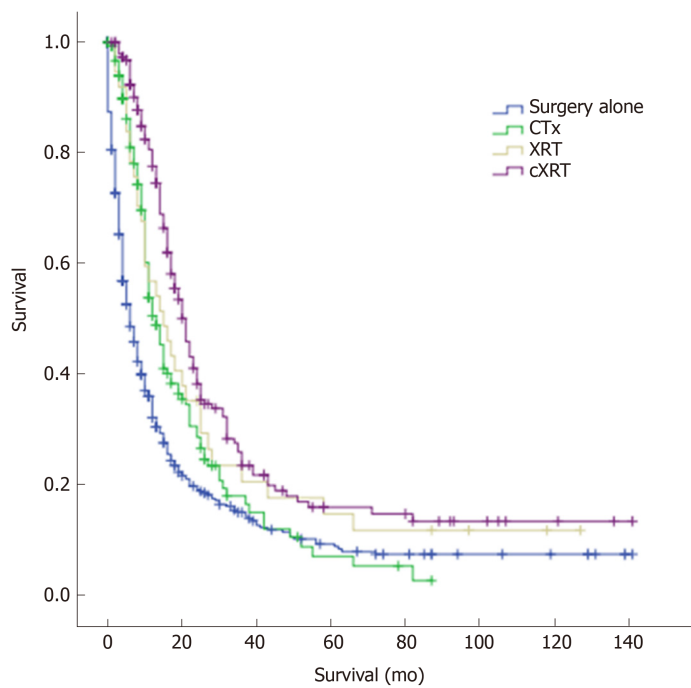
	Survival time ( <i>n</i> )	≤ 9 m ( <i>n</i> = 114)	> 9 m ( <i>n</i> = 130)
Confusion matrix ( <i>n</i> )	≤ 9 m (128)	84	44
	> 9 m (116)	30	86
Reliability (%)	≤ 9 m (128)	65.62%	34.38%
	> 9 m (116)	25.86%	74.14%
Accuracy (%)	≤ 9 m (128)	73.68%	33.85%
	> 9 m (116)	26.32%	66.15%

benefit from the addition of XRT and cXRT. For patients with node-positive disease, the model predicted survival benefit from xCRT.

**Table 3** Survival prediction table as established by T stage, N stage, radiation, and chemotherapy

T stage	N stage	Radiation	Chemotherapy	> 9 m (%)
3	0	0	0	41.35
3	0	0	1	58.29
3	0	1	0	75.42
3	0	1	1	76.62
3	1	0	0	34.59
3	1	0	1	64.08
3	1	1	0	69.71
3	1	1	1	80.71
4	0	0	0	28.50
4	0	0	1	44.14
4	0	1	0	63.44
4	0	1	1	64.95
4	1	0	0	14.85
4	1	0	1	37.03
4	1	1	0	43.14
4	1	1	1	57.97

T stage: 3 T3 stage, 4 T4 stage; N stage: 0 N0 stage, 1 N1 stage; Radiation and chemotherapy: 0 No, 1 Yes.



No. at risk	0	20	40	60	80	100	120	140
Surgery alone	436	69	31	20	13	8	6	1
CTx	156	36	10	4	2			
XRT	37	14	7	5	4	2	1	
cXRT	189	72	25	13	11	6	3	2

**Figure 3** Overall survival curve with different adjuvant therapies for patients with advanced gallbladder adenocarcinoma. CTx: Adjuvant chemotherapy; XRT: Adjuvant radiotherapy; cXRT: Adjuvant chemoradiotherapy.

## ARTICLE HIGHLIGHTS

### Research background

Gallbladder carcinoma (GBC) is the most common biliary tract cancer and the sixth most common gastrointestinal malignancy worldwide. Surgical resection is the only potentially curative treatment for GBC, and the outcome of patients with advanced disease is dismal. The

factors affecting the prognosis and role of adjuvant therapy in advanced GBC after curative resection remain unclear.

### Research motivation

In order to identify the factors affecting the prognosis and role of adjuvant therapy in advanced GBC after curative resection, the establishment of an accurate survival prediction model for patients with advanced GBC is of great significance for the selection of individualized treatments to increase the survival time. We have previously applied a Bayesian network (BN) and importance measures to identify the significant factors of survival after surgery for patients with GBC. In the present study, we applied BN to build a model to predict the survival time for patients with advanced GBC following curative resection.

### Research objectives

The objective of this study was to provide a survival prediction model and decision-making support to patients with advanced GBC after curative resection, as well as to identify the prognostic factors associated with survival and the role of adjuvant therapy.

### Research methods

Patients with curatively resected advanced gallbladder adenocarcinoma (T3 and T4) were selected from the Surveillance, Epidemiology, and End Results (SEER) database between 2004 and 2015. We constructed a survival prediction model based on the SEER database using the tree-augmented naïve Bayes algorithm, and composite importance measures were applied to rank the influence of prognostic factors on survival. The confusion matrix and receiver operating characteristic curve were used to evaluate the model accuracy.

### Research results

A total of 818 patients met the inclusion criteria. The median survival time was 9.0 mo. The accuracy of the BN model was 69.67%, and the area under the curve value for the testing dataset was 77.72%. The importance measures showed that adjuvant radiation, adjuvant chemotherapy (CTx), T stage, scope Reg Ln Sur, and radiation sequence were ranked as the top 5 prognostic factors. A survival prediction table was then established based on T stage, N stage, adjuvant radiotherapy (XRT), and CTx. The prediction model showed that the survival time (>9.0 mo) was affected by different treatments with the order of adjuvant chemoradiotherapy > adjuvant radiation > adjuvant chemotherapy > surgery alone. For patients with node-positive disease, the larger benefit predicted by the model is adjuvant chemoradiotherapy, and the results were validated by the survival analysis further.

### Research conclusions

A BN model was constructed to predict the survival time for patients with advanced GBC after curative resection from the SEER database, with a high model accuracy. The prediction model supported the role of adjuvant therapy for advanced GBC patients. For patients with node-negative disease, the model estimated the survival benefit from the addition of XRT and cXRT. For patients with node-positive disease, adjuvant chemoradiotherapy is expected to improve the survival significantly.

### Research perspectives

The BN-based survival prediction model can be used as a decision-making support tool for advanced GBC patients. We will improve the model accuracy based on more data. Large-volume, prospective, randomized, controlled clinical trials are needed to validate the prediction model in the future.

## REFERENCES

- 1 **Hundal R**, Shaffer EA. Gallbladder cancer: epidemiology and outcome. *Clin Epidemiol* 2014; **6**: 99-109 [PMID: 24634588 DOI: 10.2147/CLEP.S37357]
- 2 **Kasumova GG**, Tabatabaie O, Najarian RM, Callery MP, Ng SC, Bullock AJ, Fisher RA, Tseng JF. Surgical Management of Gallbladder Cancer: Simple Versus Extended Cholecystectomy and the Role of Adjuvant Therapy. *Ann Surg* 2017; **266**: 625-631 [PMID: 28692469 DOI: 10.1097/SLA.0000000000002385]
- 3 **Hueman MT**, Vollmer CM, Pawlik TM. Evolving treatment strategies for gallbladder cancer. *Ann Surg Oncol* 2009; **16**: 2101-2115 [PMID: 19495882 DOI: 10.1245/s10434-009-0538-x]
- 4 **Higuchi R**, Ota T, Araida T, Kajiyama H, Yazawa T, Furukawa T, Yoshikawa T, Takasaki K, Yamamoto M. Surgical approaches to advanced gallbladder cancer : a 40-year single-institution study of prognostic factors and resectability. *Ann Surg Oncol* 2014; **21**: 4308-4316 [PMID: 25023547 DOI: 10.1245/s10434-014-3885-1]
- 5 **Birnbaum DJ**, Viganò L, Russolillo N, Langella S, Ferrero A, Capussotti L. Lymph node metastases in patients undergoing surgery for a gallbladder cancer. Extension of the lymph node dissection and prognostic value of the lymph node ratio. *Ann Surg Oncol* 2015; **22**: 811-818 [PMID: 25201500 DOI: 10.1245/s10434-014-4044-4]
- 6 **Hawkins WG**, DeMatteo RP, Jarnagin WR, Ben-Porat L, Blumgart LH, Fong Y. Jaundice predicts advanced disease and early mortality in patients with gallbladder cancer. *Ann Surg Oncol* 2004; **11**: 310-315 [PMID: 14993027 DOI: 10.1245/ASO.2004.03.011]
- 7 **Mantripragada KC**, Hamid F, Shafiqat H, Olszewski AJ. Adjuvant Therapy for Resected Gallbladder Cancer: Analysis of the National Cancer Data Base. *J Natl Cancer Inst* 2016; **109** [PMID: 27707843 DOI: 10.1093/jnci/djw202]

- 8 **Mitin T**, Enestvedt CK, Jemal A, Sineshaw HM. Limited Use of Adjuvant Therapy in Patients With Resected Gallbladder Cancer Despite a Strong Association With Survival. *J Natl Cancer Inst* 2017; **109** [PMID: 28376178 DOI: 10.1093/jnci/djw324]
- 9 **Park HS**, Lim JY, Yoon DS, Park JS, Lee DK, Lee SJ, Choi HJ, Song SY, Lee WJ, Cho JY. Outcome of adjuvant therapy for gallbladder cancer. *Oncology* 2010; **79**: 168-173 [PMID: 21212704 DOI: 10.1159/000322914]
- 10 **Moons KG**, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, Vickers AJ, Ransohoff DF, Collins GS. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015; **162**: W1-73 [PMID: 25560730 DOI: 10.7326/M14-0698]
- 11 **Wang SJ**, Lemieux A, Kalpathy-Cramer J, Ord CB, Walker GV, Fuller CD, Kim JS, Thomas CR. Nomogram for predicting the benefit of adjuvant chemoradiotherapy for resected gallbladder cancer. *J Clin Oncol* 2011; **29**: 4627-4632 [PMID: 22067404 DOI: 10.1200/JCO.2010.33.8020]
- 12 **Wang SJ**, Fuller CD, Kim JS, Sittig DF, Thomas CR, Ravdin PM. Prediction model for estimating the survival benefit of adjuvant radiotherapy for gallbladder cancer. *J Clin Oncol* 2008; **26**: 2112-2117 [PMID: 18378567 DOI: 10.1200/JCO.2007.14.7934]
- 13 **Ethun CG**, Postlewait LM, Le N, Pawlik TM, Buettner S, Poultides G, Tran T, Idrees K, Isom CA, Fields RC, Jin LX, Weber SM, Salem A, Martin RC, Scoggins C, Shen P, Mogal HD, Schmidt C, Beal E, Hatzaras I, Shenoy R, Merchant N, Cardona K, Maithel SK. A Novel Pathology-Based Preoperative Risk Score to Predict Locoregional Residual and Distant Disease and Survival for Incidental Gallbladder Cancer: A 10-Institution Study from the U.S. Extrahepatic Biliary Malignancy Consortium. *Ann Surg Oncol* 2017; **24**: 1343-1350 [PMID: 27812827 DOI: 10.1245/s10434-016-5637-x]
- 14 **Pearl J**. Probabilistic reasoning in intelligent systems: Networks of plausible inference. San Francisco: Morgan Kaufmann 1988; 116-130
- 15 **Petousis P**, Han SX, Aberle D, Bui AA. Prediction of lung cancer incidence on the low-dose computed tomography arm of the National Lung Screening Trial: A dynamic Bayesian network. *Artif Intell Med* 2016; **72**: 42-55 [PMID: 27664507 DOI: 10.1016/j.artmed.2016.07.001]
- 16 **Stojadinovic A**, Bilchik A, Smith D, Eberhardt JS, Ward EB, Nissan A, Johnson EK, Protic M, Peoples GE, Avital I, Steele SR. Clinical decision support and individualized prediction of survival in colon cancer: bayesian belief network model. *Ann Surg Oncol* 2013; **20**: 161-174 [PMID: 22899001 DOI: 10.1245/s10434-012-2555-4]
- 17 **Cai ZQ**, Guo P, Si SB, Geng ZM, Chen C, Cong LL. Analysis of prognostic factors for survival after surgery for gallbladder cancer based on a Bayesian network. *Sci Rep* 2017; **7**: 293 [PMID: 28331235 DOI: 10.1038/s41598-017-00491-3]
- 18 **Lau CSM**, Zywyot A, Mahendraraj K, Chamberlain RS. Gallbladder Carcinoma in the United States: A Population Based Clinical Outcomes Study Involving 22,343 Patients from the Surveillance, Epidemiology, and End Result Database (1973-2013). *HPB Surg* 2017; **2017**: 1532835 [PMID: 28638176 DOI: 10.1155/2017/1532835]
- 19 **Kim Y**, Margonis GA, Prescott JD, Tran TB, Postlewait LM, Maithel SK, Wang TS, Evans DB, Hatzaras I, Shenoy R, Phay JE, Keplinger K, Fields RC, Jin LX, Weber SM, Salem AI, Sicklick JK, Gad S, Yopp AC, Mansour JC, Duh QY, Seiser N, Solorzano CC, Kiernan CM, Votanopoulos KI, Levine EA, Poultides GA, Pawlik TM. Nomograms to Predict Recurrence-Free and Overall Survival After Curative Resection of Adrenocortical Carcinoma. *JAMA Surg* 2016; **151**: 365-373 [PMID: 26676603 DOI: 10.1001/jamasurg.2015.4516]
- 20 **Kourou K**, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. *Comput Struct Biotechnol J* 2014; **13**: 8-17 [PMID: 25750696 DOI: 10.1016/j.csbj.2014.11.005]
- 21 **Miksad RA**. When a decision must be made: role of computer modeling in clinical cancer research. *J Clin Oncol* 2011; **29**: 4602-4604 [PMID: 22067392 DOI: 10.1200/JCO.2011.37.8604]
- 22 **Horgan AM**, Amir E, Walter T, Knox JJ. Adjuvant therapy in the treatment of biliary tract cancer: a systematic review and meta-analysis. *J Clin Oncol* 2012; **30**: 1934-1940 [PMID: 22529261 DOI: 10.1200/JCO.2011.40.5381]
- 23 **Hyder O**, Dodson RM, Sachs T, Weiss M, Mayo SC, Choti MA, Wolfgang CL, Herman JM, Pawlik TM. Impact of adjuvant external beam radiotherapy on survival in surgically resected gallbladder adenocarcinoma: a propensity score-matched Surveillance, Epidemiology, and End Results analysis. *Surgery* 2014; **155**: 85-93 [PMID: 23876364 DOI: 10.1016/j.surg.2013.06.001]
- 24 **Kim Y**, Amini N, Wilson A, Margonis GA, Ethun CG, Poultides G, Tran T, Idrees K, Isom CA, Fields RC, Krasnick B, Weber SM, Salem A, Martin RC, Scoggins C, Shen P, Mogal HD, Schmidt C, Beal E, Hatzaras I, Shenoy R, Cardona K, Maithel SK, Pawlik TM. Impact of Chemotherapy and External-Beam Radiation Therapy on Outcomes among Patients with Resected Gallbladder Cancer: A Multi-institutional Analysis. *Ann Surg Oncol* 2016; **23**: 2998-3008 [PMID: 27169772 DOI: 10.1245/s10434-016-5262-8]
- 25 **Jarnagin WR**, Ruo L, Little SA, Klimstra D, D'Angelica M, DeMatteo RP, Wagman R, Blumgart LH, Fong Y. Patterns of initial disease recurrence after resection of gallbladder carcinoma and hilar cholangiocarcinoma: implications for adjuvant therapeutic strategies. *Cancer* 2003; **98**: 1689-1700 [PMID: 14534886 DOI: 10.1002/cncr.11699]
- 26 **Hoehn RS**, Wima K, Ertel AE, Meier A, Ahmad SA, Shah SA, Abbott DE. Adjuvant Therapy for Gallbladder Cancer: an Analysis of the National Cancer Data Base. *J Gastrointest Surg* 2015; **19**: 1794-1801 [PMID: 26293376 DOI: 10.1007/s11605-015-2922-0]
- 27 **Ben-Josef E**, Guthrie KA, El-Khoueiry AB, Corless CL, Zalupski MM, Lowy AM, Thomas CR, Alberts SR, Dawson LA, Micetich KC, Thomas MB, Siegel AB, Blanke CD. SWOG S0809: A Phase II Intergroup Trial of Adjuvant Capecitabine and Gemcitabine Followed by Radiotherapy and Concurrent Capecitabine in Extrahepatic Cholangiocarcinoma and Gallbladder Carcinoma. *J Clin Oncol* 2015; **33**: 2617-2622 [PMID: 25964250 DOI: 10.1200/JCO.2014.60.2219]



Published By Baishideng Publishing Group Inc  
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA  
Telephone: +1-925-2238242  
E-mail: [bpgoffice@wjgnet.com](mailto:bpgoffice@wjgnet.com)  
Help Desk: <http://www.f6publishing.com/helpdesk>  
<http://www.wjgnet.com>

