Development and validation of a nomogram for survival benefit of lymphadenectomy in resected gallbladder cancer

Mingyu Chen[1,2#], Jian Lin[3#], Jiasheng Cao[1#], Hepan Zhu[1], Bin Zhang[1], Angela Wu[4], Xiujun Cai[1,2]

[1]Department of General Surgery, [2]Key Laboratory of Endoscopic Technique Research of Zhejiang Province, Sir Run-Run Shaw Hospital, Zhejiang University, Hangzhou 310016, China; [3]Longyou People's Hospital, Quzhou 324400, China; [4]Medicine, University of Melbourne, Melbourne, VIC, Australia

Contributions: (I) Conception and design: M Chen, J Lin, X Cai; (II) Administrative support: None; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: J Lin, H Zhu; (V) Data analysis and interpretation: B Zhang, A Wu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Dr. Xiujun Cai. Department of General Surgery, Sir Run-Run Shaw Hospital, and Key Laboratory of Endoscopic Technique Research of Zhejiang Province, Zhejiang University, No. 3 East Qingchun Road, Hangzhou 310016, China. Email: srrsh_cxj@zju.edu.cn.

Background: Due to absence of large, prospective, randomized, clinical trial data, the potential survival benefit of lymphadenectomy with different number of regional lymph nodes (LNs) remains controversial. We aim to create a predicting model to help estimate individualized potential survival benefit of lymphadenectomy with more regional LNs for patients with resected gallbladder cancer (GBC).

Methods: Patients with resected GBC were selected from the Surveillance, Epidemiology, and End Results database who were diagnosed between 2004 and 2014. Covariates included age, race, sex, grade, histological stage, tumor sizes and receipt of non-primary surgery. Two types of multivariate survival regression models were constructed and compared. The best model performance was tested by the external validation data from our hospital.

Results: A total of 1,669 patients met the inclusion criteria for this study. The lognormal survival model showed the best performance and was tested by the external validation data, including 193 patients with resected GBC from our hospital. Nomograms, which based on the accelerated failure time parametric survival model, were built to estimate individualized survival. C-index, was up to 0.754 and 0.710 in internal validation for more and less regional LNs removed, respectively. Both of internal and external calibration curves showed good agreement between predicted and observed outcomes in the 1-, 3-, and 5-year overall survival (OS).

Conclusions: A predicting model can be used as a decision model to predict which patients may obtain benefit from lymphadenectomy with more regional LNs.

Keywords: Gallbladder cancer (GBC); nomogram; lymphadenectomy; predicting model; overall survival (OS)

Introduction

Gallbladder cancer (GBC) is one of the most common biliary tract malignancies (1,2). It presents with a low annual incidence (3), poor prognosis (1,3,4) and high mortality, because of a high proportion of early lymph node (LN) metastases (5). LN status plays an important role in prognosis (6-8), i.e., positive LN status indicates a poor prognosis. Surgery remains the first line therapy for patients with resected GBC (9). There is much variation in literatures about what composes a "radical cholecystectomy" and/or an "appropriate lymphadenectomy" (10). For complete resection, extended surgical procedures,

such as major hepatectomy and adequate extensive lymphadenectomy, even common bile duct resection or pancreatoduodenectomy, are often required (11). Although lymphadenectomy enables to remove more regional LNs and facilitate accurate staging of cancer, it also increases the operative difficulty and risk. Furthermore, even after extensive lymphadenectomy, not all patients with GBC can benefit from it. Therefore, whether lymphadenectomy with more regional LNs for a risk of metastatic disease could enhance or contribute to this curative potential remains debated and unproven.

Several different LN staging/scoring systems, such as tumor-node-metastasis, LN ratio, the log odds of positive LN, have been proposed to stratify the prognosis of patients with GBC. Unfortunately, none of them focuses on this debate. Moreover, it is not easy and convenient to make decision on who need to be performed with more regional lymphadenectomy during operation, according to the results of frozen-sections alone. Due to the rarity of GBC and the lack of large-scale prospective randomized clinical trials, the actual benefit for removing more and/or less regional LNs at risk of GBC has not been well established. As a result, there is little evidence for clinicians to rely on to determine which patients will obtain benefit from more regional LNs.

The primary aim of this study was to create a decision model to estimate individualized potential survival benefit of lymphadenectomy with more and/or less regional LNs for patients with resected GBC.

Methods

Study population

Surveillance, Epidemiology and End Results (SEER) database, which covers approximately 26% of the U.S. population, provides patients' data, including patient demographics, tumor morphology, staging, treatment detail, follow-up and so on. Patients who underwent resection for GBC between 2004 and 2014 were identified in the SEER database of the National Cancer Institute. GBC was identified using the International Classification of Diseases (ICD-O-3) (C23.9) codes, and patients diagnosed at autopsy, or none of regional LNs removed were excluded. Standard patient demographic and clinicopathologic data, including size, grade, and histological stage, was collected. In addition, one to three regional LNs removed was defined as "lymphadenectomy with less regional LNs", while four or more regional LNs removed was defined as "lymphadenectomy with more regional LNs". Patients

who underwent resection for GBC from January 2007 to December 2012 at author's institution were included in the study as external validation data.

Statistical analysis

Statistical analyses were performed using the SPSS 24.0 or R software packages. The primary end point of interest in this study was overall survival (OS). Observed covariates were age, sex, race, grade, tumor size, American Joint Commotion Cancer of T stage according to 7th edition and receipt of non-primary surgery and so on. A propensity score 1:1 matching method was performed to balance observed covariates in two groups using the SPSS 24.0. By assigning propensity score weights to each patient and incorporating these weights into model construction, we can reduce inherent biases in retrospective non-randomized regression analyses. Multivariate regression survival analysis was performed to identify significant factors. Then, two survival modeling methods such as Semiparametric model (Cox proportional hazards) and accelerated failure time parametric model (lognormal) were compared using Akaike's Information Criterion. The best model was selected and tested by the internal data form SEER database and external validation data from authors' hospital using both discrimination and calibration. Discrimination was evaluated using the Harrell's concordance index (C-index). Calibration, which compares predicted with actual survival, was evaluated with a calibration curve. Except that, the analysis of subgroup (1 LNs and 2–3 LNs) from less regional LNs group was performed. When P value less than 0.05, it means significant. In addition, STROBE and TRIPOD guidelines are performed in the observational study to consult for prediction model.

Results

Patient and tumor characteristics

A total of 1,669 coming from SEER were summered in *Table 1*. There were some differences in two groups, such as patients undergoing lymphadenectomy with more regional LNs group tended to be younger, had lower histological differentiated grade, smaller tumor size, higher T-stages and percent of non-primary. Of these, after propensity score weighting applied to balance covariates in two groups, all covariates were balanced and no longer had statistically significant difference. In addition, the characteristics of 193 patients from our hospital were shown in *Table 2*.

Table 1 Patient and tumor characteristics before and after PS weighting applied to balance covariates between less and more regional LNs groups

Characteristic	Original (N=1,669)			PS weighted (N=1,012)		
	Less	More	P	Less	More	P
Age (years)			<0.001			0.116
<50	74	55		48	41	
50–60	189	113		107	89	
60–70	293	201		142	177	
70–80	312	153		139	142	
≥80	219	60		70	57	
Sex			0.527			0.733
Female	767	402		353	348	
Male	320	180		153	158	
Race			0.306			0.581
White	816	426		365	377	
Black	132	86		67	66	
Other	139	70		74	63	
Differentiated grade			0.007			0.574
Well	153	68		70	61	
Moderate	495	312		259	255	
Poor	439	202		177	190	
Tumor size (cm)	3.9±2.9	3.5±2.5	0.004	3.8±2.7	3.8±2.5	0.869
Histology			0.092			0.437
Adenocarcinoma	820	417		132	143	
Other	267	165		374	363	
AJCC T, 7th			<0.001			0.794
T1a	60	14		11	11	
T1b	92	35		19	18	
T2	518	264		122	131	
T3	388	237		100	89	
T4	29	32		5	8	
Non-primary surgery			<0.001			0.793
Yes	245	243		177	181	
No	842	339		506	325	

More, lymphadenectomy more regional LNs (4 or more regional lymph nodes removed); Less, lymphadenectomy less regional LNs (1 to 3 regional lymph nodes removed). LNs, lymph nodes; PS, propensity score.

Table 2 Patient and tumor characteristics between less and more regional LNs groups from SRRSH database

Characteristic	Less regional LNs (N=108)		More regional LNs (N=85)		P
	Number	%	Number	%	
Age (years)					0.151
<50	10	9.3	13	15.3	
50–60	42	38.9	22	25.9	
60–70	25	23.1	28	32.9	
70–80	25	23.1	20	23.5	
≥80	6	5.6	2	2.4	
Sex					0.930
Female	82	75.9	65	76.5	
Male	26	24.1	20	23.5	
Differentiated grade					0.030
Well	22	20.4	13	15.3	
Moderate	70	64.8	46	54.1	
Poor	16	14.8	26	30.6	
Tumor size (cm)	2.7±1.8	–	3.0±1.5	–	0.232
Histology					0.298
Adenocarcinoma	91	84.3	76	89.4	
Other	17	15.7	9	10.6	
AJCC T, 7th, %					<0.001
T1a	22	20.4	1	1.2	
T1b	18	16.7	5	5.9	
T2	59	54.6	65	76.5	
T3	9	8.3	14	16.5	
T4	–	–	–	–	
Non-primary surgery					0.086
Yes	11	10.2	16	18.8	
No	97	89.8	69	81.2	

LNs, lymph nodes; SRRSH, Sir Run Run Shaw Hospital.

Independent factors and two nomograms

The multivariate survival regression analysis was performed. There were four statistically significant factors for the group of less regional LNs removed, including age (P=0.001), tumor size (P<0.001), T-stages (P<0.001) and receipt of non-primary surgery (P=0.004), which were listed in *Table 3*. At the meantime, five factors for that of more regional LNs, consisting of age (P=0.020), sex (P=0.044), grade (P=0.043), tumor size (P=0.015) and T-stages (P<0.001), was identified and summered in *Table 3*. Two nomograms were built on basis of each independent factor. In order to compare the performance of survival models, the lognormal model had the lowest Akaike's Information Criterion of 9032, indicating a better overall fit than the Cox proportional hazards models (9546). According to the coefficients from this model, two nomograms (*Figure 1A,B*) were constructed to estimate the survival benefit for lymphadenectomy with less and more regional LNs, respectively. To use the nomogram, first draw a vertical line up to the top point row to assign points for each variable. Then, add up the total points and drop a vertical line from the total point row to

Table 3 Multivariate survival regression analysis results after PS weighting

Covariate	Less regional LNs			More regional LNs		
	HR	95% CI	P	HR	95% CI	P
Age (years)			0.001**			0.020*
<50	Reference			Reference		
50–60	1.165	0.715–1.896	0.540	0.726	0.411–1.281	0.269
60–70	1.365	0.847–2.199	0.201	1.121	0.675–1.863	0.660
70–80	1.480	0.925–2.370	0.102	1.106	0.660–1.854	0.702
≥80	2.386	1.433–3.973	0.001**	1.631	0.917–2.904	0.096
Sex			0.233			0.044*
Female	Reference			Reference		
Male	1.164	0.907–1.493	0.233	1.332	1.008–1.761	0.044*
Race			0.425			0.199
White	Reference			Reference		
Black	1.094	0.761–1.572	0.628	1.441	0.963–2.156	0.075
Other	0.811	0.564–1.165	0.257	1.116	0.736–1.694	0.605
Grade			0.137			0.043*
Well	Reference			Reference		
Moderate	1.118	0.754–1.658	0.577	1.495	0.887–2.521	0.131
Poor	1.400	0.923–2.122	0.113	1.876	1.101–3.197	0.021*
Tumor size	1.008	1.004–1.013	<0.001[#]	1.006	1.001–1.011	0.015*
Histology			0.097			0.077
Other	Reference			Reference		
Adenocarcinoma	1.262	0.959–1.661	0.097	1.301	1.008–1.761	0.077
AJCC T, 7th			<0.001[#]			<0.001[#]
T1a	Reference			Reference		
T1b	2.154	0.472–9.823	0.332	1.608	0.334–7.739	0.553
T2	2.605	0.635–10.678	0.183	1.846	0.445–7.659	0.398
T3	5.118	1.243–21.078	0.024*	6.392	1.539–26.556	0.011*
T4	8.159	1.822–36.541	0.006**	20.396	4.475–92.956	<0.001[#]
Non-primary surgery			0.004**			0.606
No	Reference			Reference		
Yes	0.681	0.523–0.887	0.004**	0.930	0.706–1.226	0.606

*, P<0.05; **, P<0.01; #, P<0.001. HR, hazard ratio.

obtain the 1-year OS, 3-year OS, and 5-year OS.

Performance of nomogram

Model performance was internally and externally validated for discrimination and calibration. Discrimination, as measured by the bootstrap corrected C-index, was 0.754 and 0.7103 in internal validation and 0.710 and 0.687 in external validation for more and less regional LNs, respectively. Both of internal calibration curves (*Figure 2A,B*) and the external calibration curves (*Figure 2C,D*) showed good agreement between predicted and observed outcomes in the
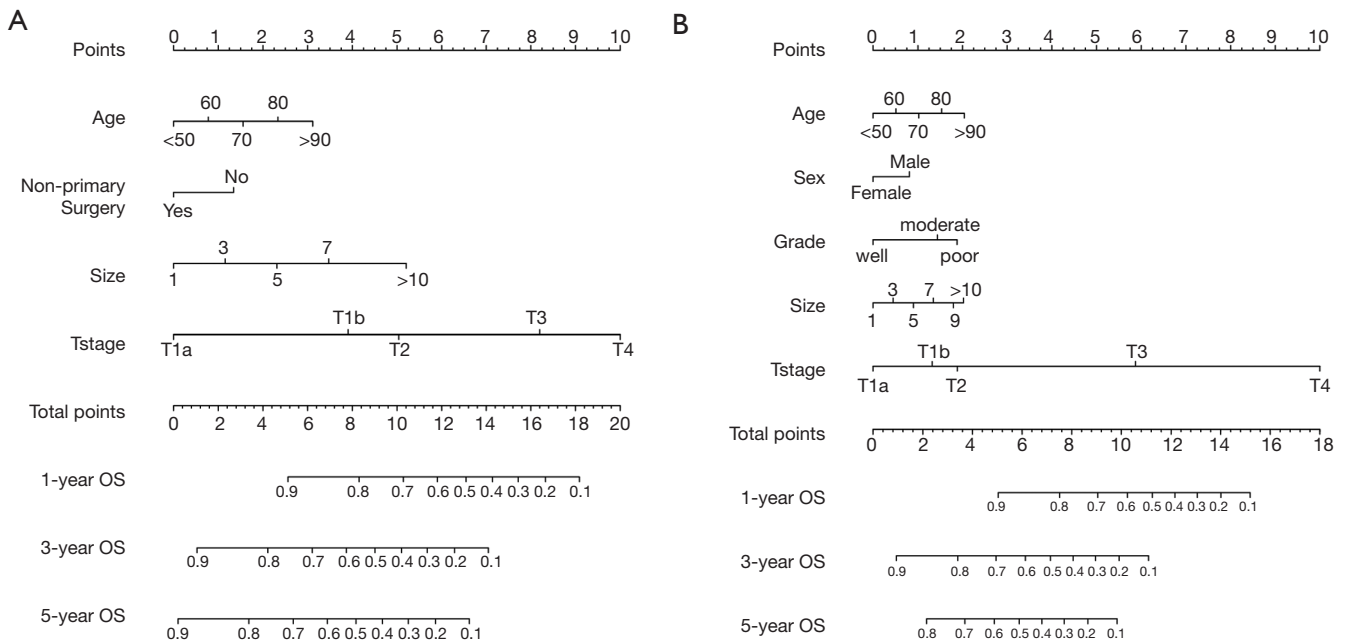
Figure 1 Nomograms for estimating benefit of lymphadenectomy for individual patient (A, less regional lymph nodes removed; B, more regional lymph nodes removed).
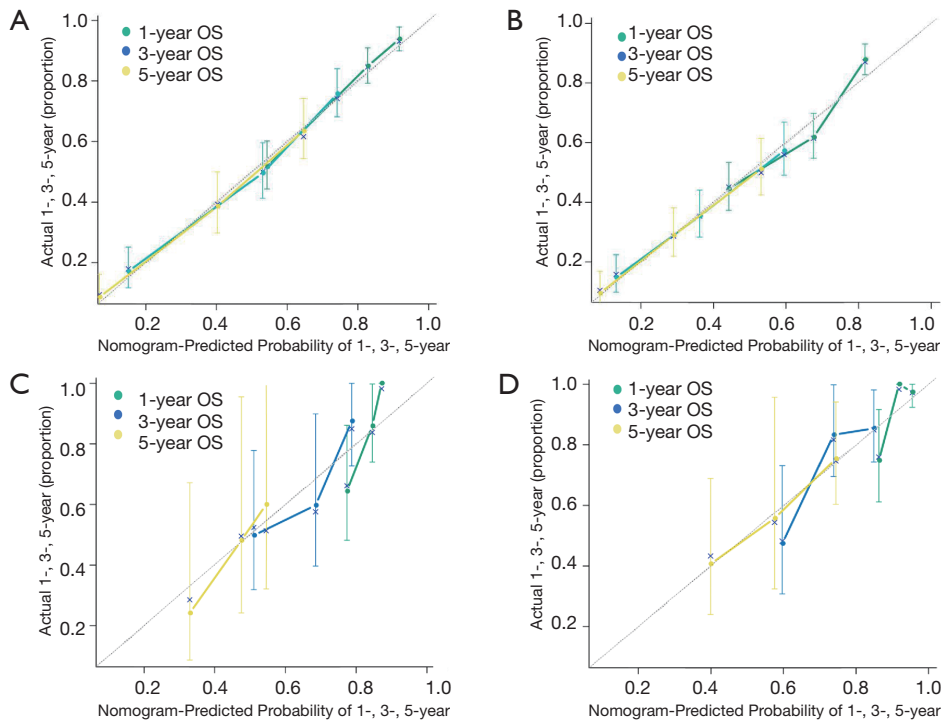

Figure 2 Internal and external calibration curve demonstrating how survival predictions from the model compare to the actual observed survival (A: internal calibration curve for 1-, 3-, 5-year OS at more regional group; B: internal calibration curve for 1-, 3-, 5-year OS at less regional group; C: external calibration curve for 1-, 3-, 5-year OS at more regional group; D: external calibration curve for 1-, 3-, 5-year OS at less regional group). OS, overall survival.
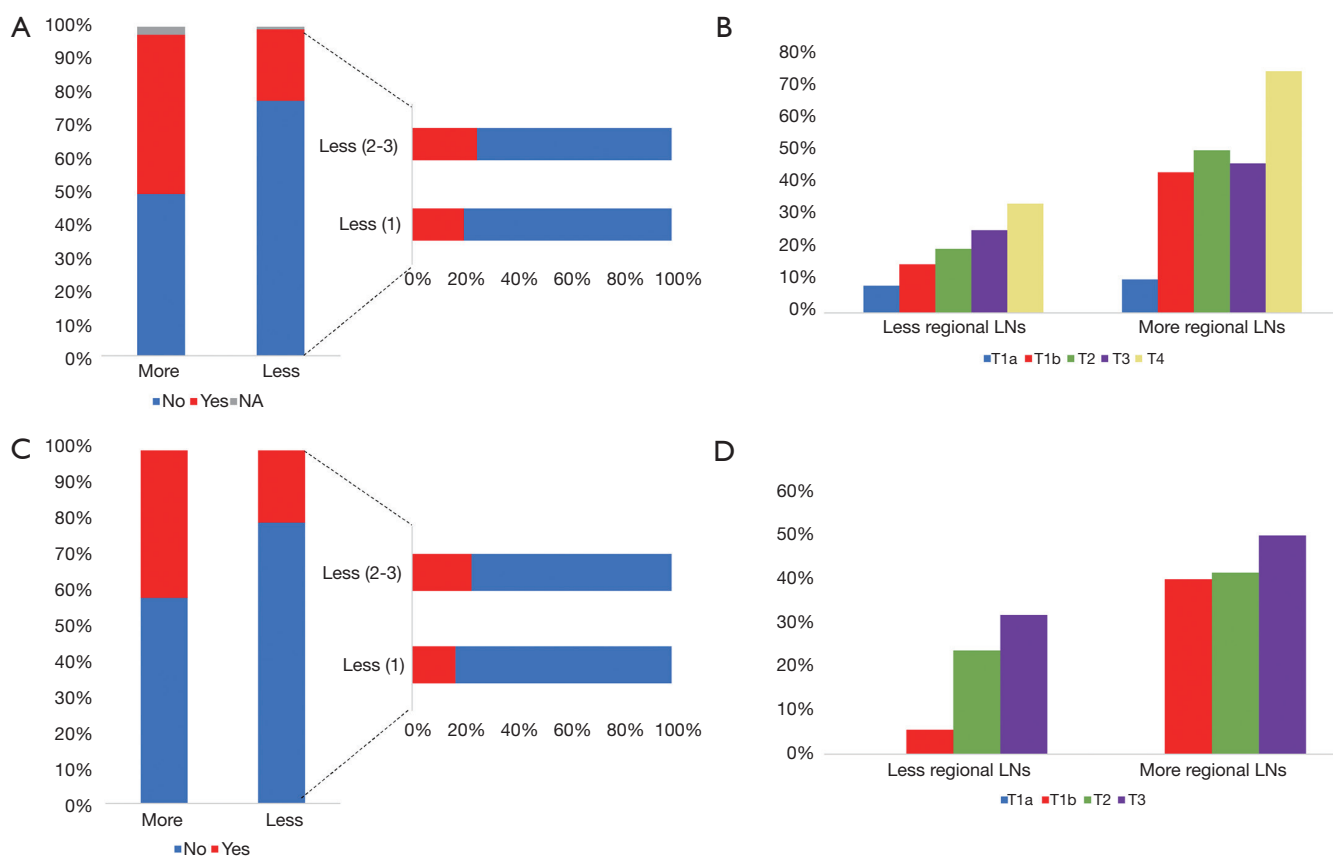
Figure 3 Difference of proportion of patients with positive LNs (A: difference between less and more regional removed in training set; B: difference in each T-stage in training set; C: difference between less and more regional removed in validation set; D: difference in each T-stage in validation set). LN, lymph node.

1-year OS, 3-year OS, and 5-year OS respectively.

Difference on the status of LNs and each T-stages

Lymphadenectomy with more regional LNs showed a higher percent of positive LNs (P<0.001), according to the data from SEER (*Figure 3A*). what's more, the percentage of positive LNs increased with higher T-staged in the less regional LNs group, while that of more regional LNs remained steady (*Figure 3B*). The same result based on hospital data were shown in *Figure 3C,D*.

Discussion

GBC is one of the most common and aggressive biliary tract malignancy (2). Because early GBCs are not always with specific symptoms, and the majority (50–70%) of them are detected as incidental findings after cholecystectomy

performed for other indications (12-14). Although the incidental gallbladder cancer (IGBC) is the most common form of GBC diagnosed today (15), many patients present with lymphatic metastases involvement (5). LN status is referred as one of the strongest prognostic factors (6). The early LNs metastasis is a characteristic of GBC (5). Fahim *et al.* (16) reported that the collecting trunks from the lymphatic plexuses in the medial and lateral wall of the gallbladder terminate in the cystic and peri-choledochal LNs and follow one of three pathways (cholecysto-mesenteric pathways, cholecysto-retropancreatic pathways and cholecysto-coeliac pathways) to converge at the para-aortic LNs between left renal vein and inferior mesenteric artery (*Figures S1,S2,S3*). Patients with LN metastasis will have a shorter survival time and 30–40% increased risk of death, comparing to patients without LNs metastasis (17,18). However, LNs status may be inaccurate without extensive lymphadenectomy. In our study, we showed that

more regional LNs removed have a higher rate of positive LNs, and a more stable accuracy in patients with different T-stage. Some studies also showed same results that it presented with LN metastases in a high proportion of patients, up to 60–80% of T3–4 tumors (7). Although the extensive lymphadenectomy was beneficial for accurately evaluating the nodal basin (19), the operative difficulty and risk was higher. Due to difficulty for surgeons to obtain the LN status directly during surgery, it is very necessary to assist clinicians in decision-making.

Cancer prediction model has been increasingly popular and important in personalized medicine (20), in which clinicians optimize the patient's therapeutic recommendations according to their specific and individual information. Recently, cancer prediction model have been used in various cancers, such as lung (21-23), breast (24-26), pancreas (27,28) and prostate cancer (29-31). Cancer prediction model usually consists of many observed covariates. Bai *et al.* (32) constructed a GBC prediction model based on covariates such as jaundice, CA19-9 and T stage to predict OS after GBC resection. However, they merged stage 0 to IIIA into one category, which made the model not accurate and specific. According to SEER database, Zhang *et al.* (33) developed a nomogram to predict prognosis in patients of GBC (M0) after surgical resection. Although they found that receipt of LN dissection was a significant variable, they didn't divide LN removal to lymphadenectomy with less regional LNs and lymphadenectomy with more regional LNs, and therefore, some biased existed in this research.

In the present study, we divided patients into less and more regional LNs removal group, meanwhile, we utilized propensity score methods which were usually used to reduce the impact of treatment selection bias, especially for non-random trails (34), to optimize the allocation of data from SEER database, and compared lognormal and Cox proportional hazards model, before we build a final survival model. Although lognormal model is not as popular as Cox proportional hazards model, it has a long history of usage in cancer survival (35) and has been shown to be a more appropriate survival model in some cancers, such as breast cancer (36), lung cancer (37), extrahepatic cholangiocarcinoma (38). Besides, Wang *et al.* (35) in current study indicated that the lognormal model also demonstrated a good fit for GBC. In this paper, we chose the accelerated failure time parametric model (lognormal), because its Akaike's Information Criterion was lower than that of semiparametric model (Cox proportional hazards).

Some factors, such as age, size and T-stage, play an important role in OS between less and more regional LNs removed group, while there were some different factors, including non-primary surgery for less regional LNs group, sex and grade for more regional LNs group. Age has a great influence on survival time in the present study as expected. Generally, the elder patients possess a poorer tolerance of stress and a damaged compensatory mechanism, and higher T-stage usually showed more aggressive of the biological behavior of malignant tumors. We found the younger and/or lower T-stage patients with GBC, the better OS, which was similar to previous studies. Interestingly, we found non-primary surgery or IGBC also a significant factor in the less regional LNs. Patients who underwent non-primary surgery usually showed lower stage, and less regional LNs removed might be enough. The other factor in more regional LNs was no difference with previous studies.

There are some limitations that need to be considered in the present study. Firstly, although the largest series of GBC cases are available form SEER, some of the known survival predictors are nearly all missing in the SEER data, such as margin status, chemoradiotherapy. Its accuracy may be affected, but this nomogram based on information which clinician can obtain before and during surgery. In addition, we used propensity score methods to reduce the impact of treatment selection bias. Therefore, postoperative treatment, such as chemoradiotherapy, immunotherapy and so on, may have little influence on the accuracy of this nomogram. Secondly, this is a retrospective study. The performance of this nomogram shows good in our hospital data, but whether it is suitable for other centers need more data to be improved and testified. Therefore, in the future, we hope get a large external data to optimize this nomogram. Finally, the details of regional LNs are missing, which increases difficulty to study which regional LNs or how much number of regional LNs are recommended to be removed. If possible, our future study will focus on this point.

In summary, we present a novel prediction model that can estimate individual survival benefit of lymphadenectomy with more and/or less regional LNs for resected GBC patients. It can be regarded as a tool to help clinician estimate which people need more regional LNs removed during surgical resection of GBC.

Acknowledgments

Development Plan Projects of Zhejiang Province (No. 2017C01018).

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

Ethical Statement: The study was approved by Sir Run Run Shaw Hospital Ethics Committee.

References

1. Yee K, Sheppard BC, Domreis J, et al. Cancers of the gallbladder and biliary ducts. Oncology (Williston Park) 2002;16:939-46, 949; discussion 949-50, 952-3, 956-7.
2. Misra S, Chaturvedi A, Misra NC, et al. Carcinoma of the gallbladder. Lancet Oncol 2003;4:167-76.
3. Randi G, Malvezzi M, Levi F, et al. Epidemiology of biliary tract cancers: an update. Ann Oncol 2009;20:146-59.
4. Ferretti S, Gafa L. Upper gastrointestinal tract cancers: oesophagus, stomach, liver, gallbladder and biliary ducts, pancreas. Epidemiol Prev 2004;28:34-42.
5. Donohue JH, Stewart AK, Menck HR. The National Cancer Data Base report on carcinoma of the gallbladder, 1989-1995. Cancer 1998;83:2618-28.
6. Tsukada K, Kurosaki I, Uchida K, et al. Lymph node spread from carcinoma of the gallbladder. Cancer 1997;80:661-7.
7. Fong Y, Wagman L, Gonen M, et al. Evidence-based gallbladder cancer staging: changing cancer staging by analysis of data from the National Cancer Database. Ann Surg 2006;243:767-71; discussion 771-4.
8. Shirai Y, Wakai T, Hatakeyama K. Radical lymph node dissection for gallbladder cancer: indications and limitations. Surg Oncol Clin N Am 2007;16:221-32.
9. Kondo S, Takada T, Miyazaki M, et al. Guidelines for the management of biliary tract and ampullary carcinomas: surgical treatment. J Hepatobiliary Pancreat Surg 2008;15:41-54.
10. Søreide K, Harrison EM, Wigmore SJ. Research gaps and unanswered questions in gallbladder cancer. HPB (Oxford) 2018;20:685-6.
11. Birnbaum DJ, Vigano L, Ferrero A, et al. Locally advanced gallbladder cancer: which patients benefit from resection? Eur J Surg Oncol 2014;40:1008-15.
12. Butte JM, Matsuo K, Gonen M, et al. Gallbladder cancer: differences in presentation, surgical treatment, and survival
in patients treated at centers in three countries. J Am Coll Surg 2011;212:50-61.
13. Ethun CG, Le N, Lopez-Aguiar AG, et al. Pathologic and Prognostic Implications of Incidental versus Nonincidental Gallbladder Cancer: A 10-Institution Study from the United States Extrahepatic Biliary Malignancy Consortium. Am Surg 2017;83:679-86.
14. Søreide K, Guest RV, Harrison EM, et al. Systematic review of management of incidental gallbladder cancer after cholecystectomy. Br J Surg 2019;106:32-45.
15. Pitt SC, Jin LX, Hall BL, et al. Incidental gallbladder cancer at cholecystectomy: when should the surgeon be suspicious? Ann Surg 2014;260:128-33.
16. Fahim RB, Mc DJ, Richards JC, et al. Carcinoma of the gallbladder: a study of its modes of spread. Ann Surg 1962;156:114-24.
17. Liu GJ, Li XH, Chen YX, et al. Radical lymph node dissection and assessment: Impact on gallbladder cancer prognosis. World J Gastroenterol 2013;19:5150-8.
18. Scheingraber S, Justinger C, Stremovskaia T, et al. The standardized surgical approach improves outcome of gallbladder cancer. World J Surg Oncol 2007;5:55.
19. Ito H, Ito K, D'Angelica M, et al. Accurate staging for gallbladder cancer: implications for surgical therapy and pathological assessment. Ann Surg 2011;254:320-5.
20. Freedman AN, Seminara D, Gail MH, et al. Cancer risk prediction models: a workshop on development, evaluation, and application. J Natl Cancer Inst 2005;97:715-23.
21. Desseroit MC, Visvikis D, Tixier F, et al. Development of a nomogram combining clinical staging with (18) F-FDG PET/CT image features in non-small-cell lung cancer stage I-III. Eur J Nucl Med Mol Imaging 2016;43:1477-85.
22. Liang W, Zhang L, Jiang G, et al. Development and validation of a nomogram for predicting survival in patients with resected non-small-cell lung cancer. J Clin Oncol 2015;33:861-9.
23. Zeng Q, Xue N, Dai D, et al. A Nomogram based on Inflammatory Factors C-Reactive Protein and Fibrinogen to Predict the Prognostic Value in Patients with Resected Non-Small Cell Lung Cancer. J Cancer 2017;8:744-53.
24. Dihge L, Bendahl PO, Ryden L. Nomograms for preoperative prediction of axillary nodal status in breast cancer. Br J Surg 2017;104:1494-505.
25. Dingemans SA, de Rooij PD, van der Vuurst de Vries RM, et al. Validation of Six Nomograms for Predicting Non-sentinel Lymph Node Metastases in a Dutch Breast Cancer Population. Ann Surg Oncol 2016;23:477-81.

26. Tsoutsou PG, Jeanneret Sozzi W, Matzinger O, et al. Nomograms predicting locoregional recurrence in the subtype era of breast cancer. J Clin Oncol 2013;31:647-8.

27. Hijioka S, Shimizu Y, Mizuno N, et al. Can long-term follow-up strategies be determined using a nomogram-based prediction model of malignancy among intraductal papillary mucinous neoplasms of the pancreas? Pancreas 2014;43:367-72.

28. Jang JY, Park T, Lee S, et al. Proposed Nomogram Predicting the Individual Risk of Malignancy in the Patients With Branch Duct Type Intraductal Papillary Mucinous Neoplasms of the Pancreas. Ann Surg 2017;266:1062-8.

29. Brockman JA, Alanee S, Vickers AJ, et al. Nomogram Predicting Prostate Cancer-specific Mortality for Men with Biochemical Recurrence After Radical Prostatectomy. Eur Urol 2015;67:1160-7.

30. Hirasawa Y, Nakashima J, Sugihara T, et al. Development of a Nomogram for Predicting Severe Neutropenia Associated With Docetaxel-Based Chemotherapy in Patients With Castration-Resistant Prostate Cancer. Clin Genitourin Cancer 2017;15:176-81.

31. Lughezzani G, Lazzeri M, Haese A, et al. Multicenter European external validation of a prostate health index-based nomogram for predicting prostate cancer at extended biopsy. Eur Urol 2014;66:906-12.

32. Bai Y, Liu ZS, Xiong JP, et al. Nomogram to predict overall survival after gallbladder cancer resection in China. World J Gastroenterol 2018;24:5167-78.

33. Zhang W, Hong HJ, Chen YL. Establishment of a Gallbladder Cancer-Specific Survival Model to Predict Prognosis in Non-metastatic Gallbladder Cancer Patients After Surgical Resection. Dig Dis Sci 2018;63:2251-8.

34. D'Agostino RB, Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Stat Med 1998;17:2265-81.

35. Wang SJ, Lemieux A, Kalpathy-Cramer J, et al. Nomogram for predicting the benefit of adjuvant chemoradiotherapy for resected gallbladder cancer. J Clin Oncol 2011;29:4627-32.

36. Chapman JA, Lickley HL, Trudeau ME, et al. Ascertaining prognosis for breast cancer in node-negative patients with innovative survival analysis. Breast J 2006;12:37-47.

37. Tai P, Chapman JA, Yu E, et al. Disease-specific survival for limited-stage small-cell lung cancer affected by statistical method of assessment. BMC Cancer 2007;7:31.

38. Fuller CD, Wang SJ, Choi M, et al. Multimodality therapy for locoregional extrahepatic cholangiocarcinoma: a population-based analysis. Cancer 2009;115:5175-83.

Chen *et al*.

Liver

Stomach

Gallbladder

12 h

12 c   12 a

8   9

14

12 b

12 p   16 b1   Pancreas

13   Spleen

Duodenum

Gallbladder

12c   12b

14   12p, 13   12a,8,9

16b1

Thoracic duct

8-Common hepatic artery LN
9-Coeliac LN
12a-Proper hepatic artery LN
12b-Pericholedochal LN
12c-Cystic LN

12p-Retroportai LN
12h-Hilar LN
13-Posterior superior pancreaticoduodenal LN
14-Superior mesenteric artery LN
16b1-Para-aortic LN
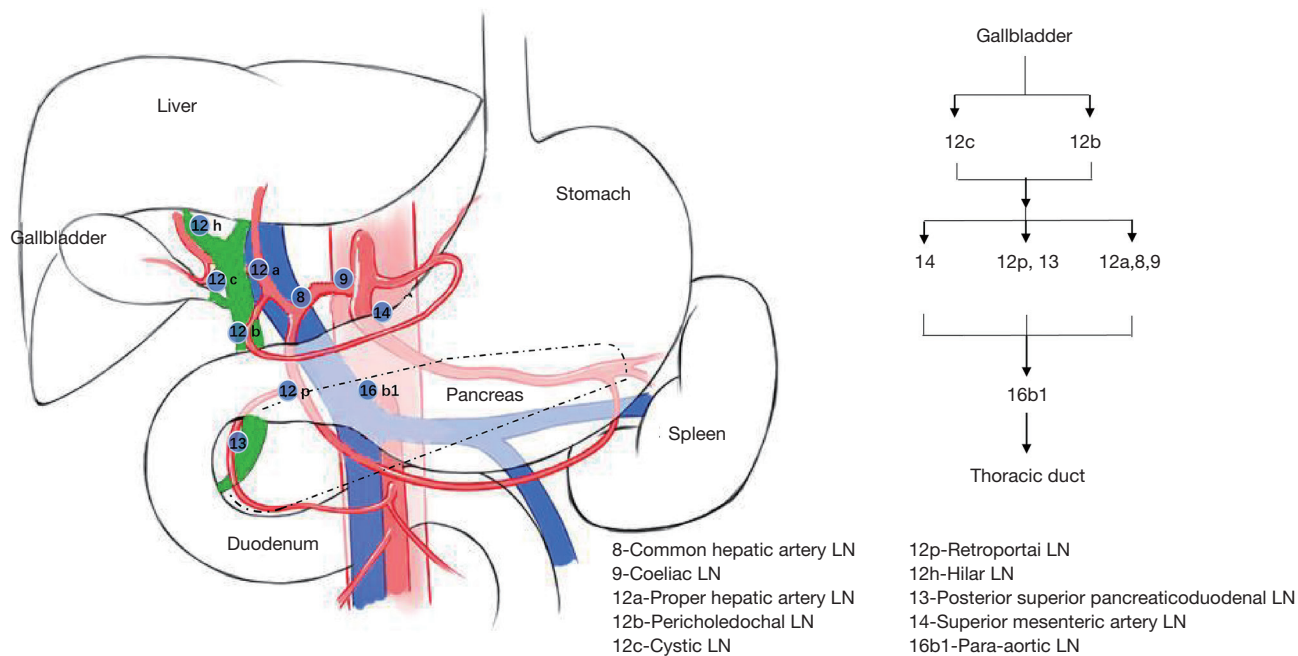
Figure S1 Pathway of lymphatic metastasis related GBC. GBC, gallbladder cancer.

Section/Topic	Item		Checklist Item	Page
Title and abstract				
Title	1	D;V	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	1
Abstract	2	D;V	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	2
Introduction				
Background and objectives	3a	D;V	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	3
	3b	D;V	Specify the objectives, including whether the study describes the development or validation of the model or both.	3
Methods				
Source of data	4a	D;V	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	4; 4
	4b	D;V	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	4
Participants	5a	D;V	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	4
	5b	D;V	Describe eligibility criteria for participants.	4
	5c	D;V	Give details of treatments received, if relevant.	4
Outcome	6a	D;V	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	4
	6b	D;V	Report any actions to blind assessment of the outcome to be predicted.	NA
Predictors	7a	D;V	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	4; 5
	7b	D;V	Report any actions to blind assessment of predictors for the outcome and other predictors.	NA
Sample size	8	D;V	Explain how the study size was arrived at.	4
Missing data	9	D;V	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	4
Statistical analysis methods	10a	D	Describe how predictors were handled in the analyses.	4; 5
	10b	D	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	4; 5
	10c	V	For validation, describe how the predictions were calculated.	4; 5
	10d	D;V	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	4; 5
	10e	V	Describe any model updating (e.g., recalibration) arising from the validation, if done.	NA
Risk groups	11	D;V	Provide details on how risk groups were created, if done.	4;5
Development *vs.* validation	12	V	For validation, identify any differences from the development data in setting, eligibility criteria, outcome, and predictors.	4; 5
Results				
Participants	13a	D;V	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	5
	13b	D;V	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	5
	13c	V	For validation, show a comparison with the development data of the distribution of important variables (demographics, predictors and outcome).	5; 6
Model development	14a	D	Specify the number of participants and outcome events in each analysis.	5
	14b	D	If done, report the unadjusted association between each candidate predictor and outcome.	NA
Model specification	15a	D	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	5
	15b	D	Explain how to the use the prediction model.	5
Model performance	16	D;V	Report performance measures (with CIs) for the prediction model.	6
Model-updating	17	V	If done, report the results from any model updating (i.e., model specification, model performance).	NA
Discussion				
Limitations	18	D;V	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	8
Interpretation	19a	V	For validation, discuss the results with reference to performance in the development data, and any other validation data.	7; 8
	19b	D;V	Give an overall interpretation of the results, considering objectives, limitations, results from similar studies, and other relevant evidence.	7; 8
Implications	20	D;V	Discuss the potential clinical use of the model and implications for future research.	8
Other information				
Supplementary information	21	D;V	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	8; 9
Funding	22	D;V	Give the source of funding and the role of the funders for the present study.	9

Items relevant only to the development of a prediction model are denoted by D, items relating solely to a validation of a prediction model are denoted by V, and items relating to both are denoted D;V. We recommend using the TRIPOD Checklist in conjunction with the TRIPOD Explanation and Elaboration document.

Figure S2 TRIPOD checklist: prediction model development and validation.

	Item No	Recommendation	Page No.
Title and abstract	1	(a) Indicate the study's design with a commonly used term in the title or the abstract	1, 2
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	2
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	3
Objectives	3	State specific objectives, including any prespecified hypotheses	3
Methods			
Study design	4	Present key elements of study design early in the paper	4
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	4
Participants	6	(a) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	4
		(b) For matched studies, give matching criteria and number of exposed and unexposed	4
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	4
Data sources/measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	4
Bias	9	Describe any efforts to address potential sources of bias	4
Study size	10	Explain how the study size was arrived at	4
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	4
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	4, 5
		(b) Describe any methods used to examine subgroups and interactions	5
		(c) Explain how missing data were addressed	4
		(d) If applicable, explain how loss to follow-up was addressed	NA
		(e) Describe any sensitivity analyses	5
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—e.g., numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	5
		(b) Give reasons for non-participation at each stage	5
		(c) Consider use of a flow diagram	NA
Descriptive data	14*	(a) Give characteristics of study participants (e.g., demographic, clinical, social) and information on exposures and potential confounders	5
		(b) Indicate number of participants with missing data for each variable of interest	5
		(c) Summarise follow-up time (e.g., average and total amount)	NA
Outcome data	15*	Report numbers of outcome events or summary measures over time	5
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included	5
		(b) Report category boundaries when continuous variables were categorized	NA
		(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	NA
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	6
Discussion			
Key results	18	Summarise key results with reference to study objectives	6, 7, 8
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	8
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	6,7,8
Generalisability	21	Discuss the generalisability (external validity) of the study results	8
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	9

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

Figure S3 STROBE statement—checklist of items that should be included in reports of cohort studies.