



Research Paper

A New Metastatic Lymph Node Classification-based Survival Predicting Model in Patients With Small Bowel Adenocarcinoma: A Derivation and Validation Study

Shan Wu¹, Jin-Nan Chen¹, Qing-Wei Zhang¹, Chao-Tao Tang¹, Xin-Tian Zhang, Ming-Yu Tang, Xiao-Bo Li^{*}, Zhi-Zheng Ge^{*}

Division of Gastroenterology and Hepatology, Key Laboratory of Gastroenterology and Hepatology, Ministry of Health, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai Institute of Digestive Disease, Shanghai, China



ARTICLE INFO

Article history:

Received 11 March 2018

Received in revised form 23 April 2018

Accepted 17 May 2018

Available online 13 June 2018

Keywords:

Small bowel carcinoma
Metastatic lymph node
Survival predicting model
TNM staging

ABSTRACT

Background: Current methods of lymph node (LN) staging are controversial in predicting the survival of SBA. We aimed to develop an alternative LN-classification-based nomogram to individualize SBA prognosis.

Methods: Based on the data from the Surveillance, Epidemiology, and End Results (SEER) database of patients diagnosed with SBA between 2004 and 2014, we identified the cut-off points for the number of LNs examined and the number found to be metastatic using the K-adaptive partitioning (KAPS) algorithm. Using metastatic LNs, a nomogram predicting the survival of SBA was derived, internally and externally validated, and measured by calibration curve, C-index, and decision curve analysis (DCA), and compared to the 8th TNM stage.

Results: A total of 1516 patients were included. The cut-off of 17 was the optimal examined LN number. For metastatic LN numbers, the cut-off points were 0, 2, and 8. The C-index for the nomogram was higher than the 8th TNM staging (internal: 0.734; 95% CI, 0.693 to 0.775 vs. 0.677; 95% CI, 0.652 to 0.702, $P < 0.001$; external: 0.715; 95% CI, 0.674 to 0.756 vs. 0.648; 95% CI, 0.602 to 0.693, $P < 0.001$). Also, the nomogram showed good calibration in internal and external validation and larger net benefit than TNM staging.

Conclusion: We modified current N staging into a 4-level staging system based on the number of metastatic LNs: N0, no LN metastasis; N1, 1–2 metastatic LNs; N2, 3–8 metastatic LNs, and N3, >8 metastatic LNs and set the least examined LN number to 17. A nomogram based on this staging showed great clinical usability than TNM staging for predicting the survival of SBA patients.

© 2018 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Small bowel tumors are one of the major causes of obscure gastrointestinal bleeding [9]. Small bowel adenocarcinomas (SBA) are the third most common histology of small bowel tumors [23]. Advancements in enteroscopy, capsule endoscopy, and cross-sectional imaging techniques [22] have allowed patients to be diagnosed earlier and more

accurately. Although SBA is rare in the gastrointestinal tract, it has a poorer stage-stratified, cancer-specific survival than colon cancer [11,19,26]. The current American Joint Committee on Cancer TNM staging system (TNM staging) of SBAs and the number of examined lymph nodes (LNs) remain controversial in predicting survival [18,20]. Meanwhile, other independent factors, such as age, grade, and tumor size, could also affect the survival significantly [3].

The effect as assessed by the number of LNs was not found to have a linear relationship to survival, and the cut-off points were essential. Traditional statistical methods can only divide cases into groups using artificial cut-off points to evaluate the difference in survival. The K-adaptive partitioning (KAPS) algorithm [8] is a useful tool for obtaining heterogeneous subgroups by survival and finding the best cut-off points by evaluating potential multi-splits.

For this reasons, we aimed to determine the optimal number of examined LNs and alternative staging of metastatic LN number through

Abbreviation: SBA, Small bowel adenocarcinoma; LN, lymph nodes; SEER, Surveillance, Epidemiology and End Results; KAPS, K-adaptive partitioning; DCA, decision curve analysis; CSS, cause-specific survival; HR, hazard ratio; C-index, concordance index; tdROC, time-dependent receiver operating characteristic; KM-weight, censoring weighting estimators; OS, overall survival; AUC, Area under curve; AJCC, American Joint Committee on Cancer.

^{*} Corresponding authors.

E-mail addresses: lx_b_1969@163.com, (X.-B. Li), zhizhengge@aliyun.com (Z.-Z. Ge).

¹ Contributed equally to this work.

the population-based Surveillance, Epidemiology, and End Results (SEER) database by the KAPS algorithm. We here developed and validated a nomogram based on this LN staging for predicting survival for SBA patients.

2. Methods

2.1. Study Design and Data Collection

This study was based on data from the SEER 18 database which includes incidence and survival data from multiple population-based cancer registries [13]. We initially analyzed 6440 patients over 18 years old who were diagnosed SBA between 2004 and 2014. To explore the pathogenesis and influencing factors for survival of SBA, we excluded patients with distant metastatic tumors and those who survived <3 months. Since we used the pathology of lymph node as the gold standard, we excluded patients who had not undergone surgery or for whom no detailed pathology was available, and those with unknown lymph node examination or an unknown number of metastatic lymph nodes. For further comparison of the feasibility of nomogram with the 8th TNM staging (Amin) [1], we excluded patients with unknown grade of tumor, with unknown T stage, with unknown N stage, and with unknown M stage. We also excluded patients without known prognostic characteristics, including race, tumor size and location. We then collected the clinicopathologic variables from the SEER 18 database, including age, gender, race, and location of tumor, TNM staging (8th TNM staging, shown in Appendix 1), grade of tumor, histologic grade, number of lymph nodes examined, number of positive lymph nodes, tumor size, and months survived.

2.2. Outcomes

The main outcome was to evaluate the effect of lymph nodes in SBA patients, including the optimal number of lymph nodes examined and alternative staging of metastatic lymph nodes with cause-specific survival (CSS) because SBA was the main endpoint. We further constructed a survival prediction model based on this metastatic lymph nodes staging for SBA patients and validated it by comparing it to TNM staging.

2.3. Statistical Analysis

Lymph nodes were evaluated in all of the participants. To establish the optimal number of examined lymph nodes, we use the modified KAPS algorithm described by Eo et al. [8] to categorize all cases into two groups to find the optimal set of cut-off points. We also used this algorithm to evaluate multi-group split points of metastatic lymph nodes and selected the optimal number of subgroups. Survival was compared between the subgroups using Kaplan-Meier survival curves and Cox regression analysis. A univariate and multivariate Cox regression model was used to calculate the hazard ratio (HR) and the adjusted HR of the alternative examined lymph nodes and metastatic lymph nodes for survival of SBA after adjusting for age, T stage, M stage, grade, histology, and tumor size.

We used the univariate and multivariate logistic regression model to calculate odds ratios (OR) of factors influencing further LN harvesting. In construction of the survival predicting model, we divided participants into two groups. The internal cohort included patients diagnosed with SBA between 2004 and 2010, while the external validation cohort consisted of participants diagnosed between 2011 and 2014. The model was selected and constructed using the internal cohort by

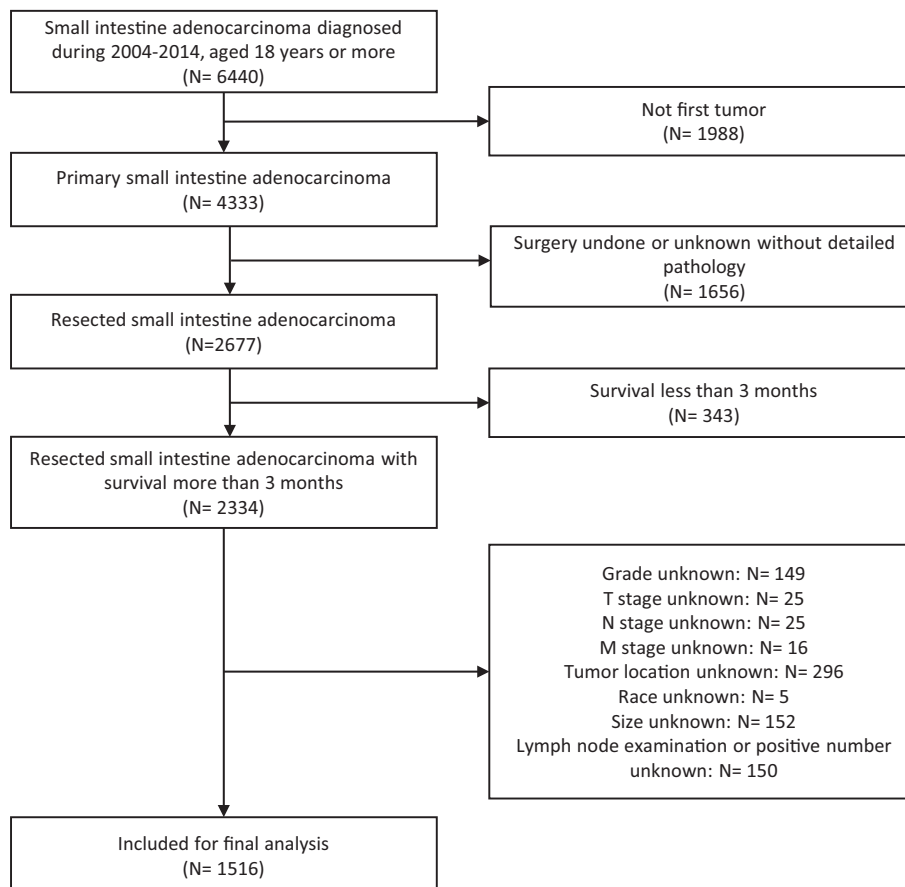


Fig. 1. Flowchart of patient selection for this study.

backward Cox analysis using AIC selection criteria, where the best model was selected with the least AIC [25]. The model was presented as nomogram and compared using the TNM staging using Harrell's concordance index (C-index) [15]. The nomogram was first internally validated using a bootstrap method and then externally validated in the independent cohorts. Time-dependent receiver operating characteristic (tdROC) curves were estimated for each cohort by inverse probability of censoring weighting estimators (KM-weight) at 1, 3, and 5 years for survival testing with the nomogram and traditional TNM staging of the specificity [4]. We also performed 1-, 3-, and 5-year CSS calibration of the nomogram by comparing the predicted survival to the observed survival in the two cohorts. Clinical usefulness and net benefit of the predictive models were estimated with decision curve analysis (DCA) and compared to traditional TNM staging throughout the whole cohort. All of the analyses were performed using R version 3.3.3.

3. Results

3.1. Patient Characteristics

From 2004 to 2014, 6440 adults were diagnosed with SBA and 1988 patients were excluded because it was not their first diagnosed tumor. Out of 4333 primary SBA patients, 1656 cases did not undergo the surgery or had no detailed pathology of the surgery. For further analysis of the relationship between LN and survival, we eliminated 343 patients who survived <3 months. In addition, patients who had missing data on any of the collected variables, including grade ($n = 149$), T stage ($n = 25$), N stage ($n = 25$), M stage ($n = 16$), tumor location ($n = 296$), race ($n = 5$), size ($n = 152$), and metastatic LN numbers ($n = 150$). Finally, 1516 cases (55% men and 45% women, mean age (\pm SD) 63 ± 53.73) were included for final analysis (Fig. 1). Baseline characteristics of participants are shown in Table 1. Among participants, 772 tumors were located in the duodenum, 396 in the jejunum, 325 in the ileum, and 23 in unspecified parts of the small intestine. There were 927 (61%) patients with Grade 1 or Grade 2 tumors and 589 (39%) with Grade 3 or 4. The actual median number of nodes harvested was 12 in the overall cohort (P25:6; P75:18), 11 in the training cohort (P25:6; P75:17), and 13 in the validation cohort (P25:8; P75:19). The average number of months of follow-up was 25 in overall cohort, 40 months in the internal cohort, and 16 months in the validation cohort. A detailed analysis of both cohorts is presented in Appendix 2.

3.2. Grouping of LNs in SBA Patients

After dividing number of examined LNs into two groups using the KAPS algorithm, we identified the optimal set of cut-off number as 17. Kaplan-Meier survival analysis performed according to the number of examined lymph nodes showed significant differences between these 2 groups ($P = 0.0021$, Fig. 2A). We then used cross-validation to obtain the best grouping of metastatic LN numbers. The group with 3 cut-off points was finally selected as the optimal set for predicting survival in SBA. The final groups selected were N0 (no LN metastasis), N1 (1–2 metastatic LNs), N2 (3–8 metastatic LNs), and N3 (>8 metastatic LNs). The difference in survival among the 4 groups was statistically significant ($P < 0.001$, Fig. 2B).

3.3. Factors that Influenced the Number of LNs Harvested

Many factors were found to influence number of LN harvest. The univariate and multivariate logistic regression was performed to explore potential factors that might be associated with the harvesting of >16 LN. As shown in the Table 3, results showed greater patient age and tumor location in the jejunum tended to be associated with less LN harvesting and more invasive stage (T2/3/4, M1). Grade 3/4 and larger tumor size tended to be associated with the harvesting of more LNs in the univariate logistic model. The multivariate logistic model,

Table 1
Characteristics of patients with small intestine adenocarcinoma.

Characteristics	Total 1516 (100%)	2004–2010 908 (60%)	2011–2014 608 (40%)	P value
Age (years)				
Mean (SD)	63 (53, 73)	63 (53, 73)	63 (54, 72)	0.625
<50	260 (17%)	152 (17%)	108 (18%)	0.125
50–75	995 (66%)	585 (66%)	410 (67%)	
>75	261 (17%)	171 (19%)	90 (15%)	
Gender				0.375
Male	828 (55%)	487 (54%)	341 (56%)	
Female	688 (45%)	421 (46%)	267 (44%)	
Race				0.261
White	1151 (76%)	701 (77%)	450 (74%)	
Black	260 (17%)	144 (16%)	116 (19%)	
Others	105 (7%)	63 (7%)	42 (7%)	
Location				0.168
Duodenum	772 (51%)	451 (50%)	321 (53%)	
Jejunum	396 (26%)	255 (28%)	141 (23%)	
Ileum	325 (21%)	187 (21%)	138 (23%)	
Small intestine, not specified	23 (2%)	15 (2%)	8 (1%)	
8th TNM stage				0.084
I	107 (7%)	61 (7%)	46 (8%)	
IIA	305 (20%)	192 (21%)	113 (19%)	
IIB	200 (13%)	104 (11%)	96 (16%)	
IIIA	413 (27%)	256 (28%)	157 (26%)	
IIIB	263 (17%)	150 (17%)	113 (19%)	
IV	228 (15%)	145 (16%)	83 (14%)	
8th T stage				0.062
T1	47 (2%)	31 (3%)	16 (2%)	
T2	99 (7%)	52 (6%)	47 (8%)	
T3	679 (45%)	429 (47%)	250 (41%)	
T4	691 (46%)	396 (44%)	295 (49%)	
8th N stage				0.099
N0	660 (44%)	387 (43%)	273 (45%)	
N1	498 (33%)	317 (35%)	181 (30%)	
N2	358 (24%)	204 (22%)	154 (25%)	
8th M stage				0.244
M0	1288 (85%)	763 (84%)	525 (86%)	
M1	228 (15%)	145 (16%)	83 (14%)	
Grade				0.057
G1/G2	927 (61%)	537 (59%)	390 (64%)	
G3/G4	589 (39%)	371 (41%)	218 (36%)	
Histology				0.390
Conventional adenocarcinoma	1330 (88%)	793 (87%)	537 (88%)	
Mucinous adenocarcinoma	128 (8%)	83 (9%)	45 (7%)	
Signet ring cell carcinoma	58 (4%)	32 (4%)	26 (4%)	
Examined lymph nodes				<0.001
≤16	1012 (67%)	641 (71%)	371 (61%)	
>16	504 (33%)	267 (29%)	237 (39%)	
Positive lymph nodes	1 (0, 3)	1 (0, 3)	1 (0, 4)	0.735
Tumor size				0.052
≤5 cm	1012 (67%)	641 (71%)	371 (61%)	
>5 cm	504 (33%)	267 (29%)	237 (39%)	
Median follow-up time (months)	25.0	40.0	16.0	<0.001

however, indicated that older age and tumor location in the jejunum were associated with the harvesting of fewer LN and more invasive stage (T2/3/4, M1), grade 3/4, and larger tumor size with harvesting of more.

3.4. SBA Survival Prediction Model

In the uni- and multi-variate Cox analysis model of clinical characteristics for prognosis of SBA, age > 75 years, T stage 4, M stage 1, Grade 3/4, number of examined LNs > 16, higher number of metastatic LNs, and tumor size > 5 cm were associated with the poorer prognosis. The number of examined lymph nodes was not included because the AIC became larger after this number was included in the nomogram,

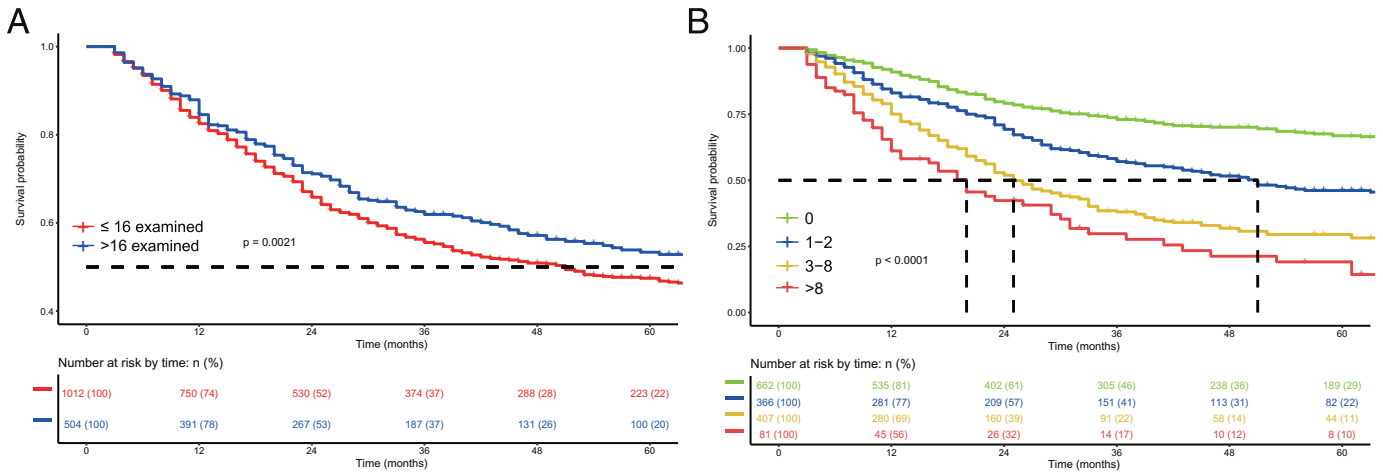


Fig. 2. Kaplan-Meier survival analysis according to the number of examined lymph nodes (A) and the number of metastatic lymph nodes (B), respectively.

and we found that was not for the model. The detailed results of Cox analysis are listed in Table 2.

Survival predicting model of nomogram was established based on selected prognostic factors (Fig. 3). The nomogram showed that T stage contributed the most to prognosis, followed by age, distant metastasis, and number of metastases. Tumor size and grade had a modest effect on survival. Each subtype of the variables was assigned a score. A straight line can be drawn down at each time point on the total point scale to determine the estimated probability of survival, according to the total number of points. For each predictor, read the points assigned on the 0–10 scale at the top and then add these points. Find the number on the “Total Points” scale and then read the corresponding predictions of 1-, 3-, and 5-year risk.

3.5. Validation and Calibration of the Nomogram

The C-indexes for the established nomogram to predict overall survival (OS) (internal: 0.734; 95% CI, 0.693 to 0.775; bootstrap corrected: 0.727) were both significantly higher than those of the 8th TNM staging (internal: 0.677; 95% CI, 0.652 to 0.702; bootstrap corrected: 0.677, $P < 0.001$) in the internal validation. External validation also showed superiority of the nomogram (0.715; 95% CI, 0.674 to 0.756) to the 8th TNM staging (0.654; 95% CI, 0.611 to 0.697, $P < 0.001$). Further analysis and comparison to subgrouping of TNM staging showed the same difference in the 2 cohorts (internal: 0.711; 95% CI, 0.684 to 0.737; bootstrap corrected: 0.710, $P < 0.001$; external: 0.656; 95% CI, 0.608 to 0.704; bootstrap corrected: 0.655, $P < 0.001$). In the analysis of specificity, the nomogram performed better than traditional TNM staging in both internal cohort (1-year AUC: 78.96 vs. 71.76, 3-year AUC: 79.32 vs. 72.25, 5-year AUC: 80.13 vs. 73.31, $P < 0.001$, Fig. 4A) and external cohort (1-year AUC: 74.92 vs. 67.25, 3-year AUC: 79.13 vs. 67.71, 5-year AUC: 81.48 vs. 67.17, $P < 0.001$, Fig. 5A) for 1-year, 3-year, and 5-year CSS. For 1-year, 3-year, and 5-year CSS, good agreement was observed between the actual observation and nomogram prediction in the internal cohort (Fig. 4B) and external cohort (Fig. 5B), as indicated by the calibration plots. When compared with the subgrouping of 8th TNM staging, the nomogram also showed greater specificity in both internal cohort (1-year AUC: 78.96 vs. 74.9, 3-year AUC: 79.32 vs. 77.04, 5-year AUC: 80.13 vs. 76.83, $P < 0.001$, Appendix 3) and external cohort (1-year AUC, 74.92 vs. 69.8, 3-year AUC: 79.13 vs. 69.34, 5-year AUC: 81.48 vs. 67.83, $P < 0.001$, Appendix 4) for 1-year, 3-year, and 5-year CSS.

DCA was used to compare the clinical usability of the nomogram to that of traditional TNM staging. Based on a continuum of potential thresholds for death (x axis) and the net benefit of using the model to risk-stratify patients (y axis) relative to assuming all patients will

survive, the DCA graphically presented that the nomogram was better than traditional TNM staging (Fig. 6) and subgrouping TNM staging (Appendix 5) under clinical conditions. Compared with traditional TNM staging, the nomogram showed a larger net benefit across the range of death risk in the analysis.

4. Discussion

Our present investigation shows that the number of LNs examined and LNs found to be metastatic to be inversely correlated with the survival of SBA patients, so we altered the number of LNs examined to 17 and introduced an alternative metastatic LN staging. This LN staging divided patients into 4 groups by the number of metastatic LNs: N0, no LN metastasis; N1, 1–2 metastatic LNs; N2, 3–8 metastatic LNs, and N3, >8 metastatic LNs. We further developed and validated, internally and externally, a nomogram based on the alternative metastatic LN staging for predicting survival in patients with SBA. This nomogram is based on six variables: grade, distant metastasis (M stage), tumor size, invasion (T stage), age, and number of metastatic LNs. This nomogram produced better and more accurate predictions in internal and external cohorts than the traditional 8th TNM staging method and showed better clinical usefulness throughout the survival as assessed by DCA.

It was noteworthy that the number of lymph nodes examined was an important factor in many gastrointestinal cancers [10,16,21], and the connection between survival and the number of lymph nodes examined was supported by some studies [18,20]. In fact, owing to inadequate lymph-node sampling, the advanced stage of SBA was not accounted for in the earlier studies of LNs in SBA [5,12,22]. Either ≥ 8 or ≥ 10 lymph nodes was identified as the optimal number of assessed lymph nodes in some studies that used the SEER database [18,20]. However, traditional statistical methods could only split cases into groups using artificial cut-off points to evaluate the difference in survival. In the present study, we used the KAPS algorithm to choose the best cut-off point for the number of LN examined and found it to be 17. This algorithm could be used to create heterogeneous subgroups by survival and we found the best cut-off points by evaluating potential multi-splits. Kaplan-Meier survival analysis grouped by 17 examined lymph nodes showed good distinction for predicting survival. The reason for this phenomenon may be that potential metastasized lymph nodes would be removed during the wider sampling of lymph nodes. However, during complete sampling of lymph nodes, the patients would be treated with appropriate adjuvant therapies based on more precise staging.

Many factors were found to influence number of LN harvested, including the size of the specimen (surgical radicality), host immunology, tumor biology, and diligence of the pathologist. In this study, we found a

Table 2
Univariate and multivariate Cox analysis of clinical characteristics for prognosis of small intestine adenocarcinoma for 5 year CSS.

Characteristics	Univariate analysis		Multivariate analysis	
	HR (95%CI)	P value	HR (95%CI)	P value
Age				
<50	Reference		Reference	
50–75	1.20 (0.96–1.51)	0.102	1.26 (1.01–1.58)	0.043
>75	1.83 (1.41–2.38)	<0.001	2.19 (1.67–2.86)	<0.001
Gender				
Male	Reference			
Female	0.91 (0.78–1.06)	0.240		
Race				
White	Reference			
Black	0.98 (0.79–1.21)	0.845		
Others	1.14 (0.84–1.54)	0.392		
Location				
Duodenum	Reference			
Jejunum	0.89 (0.74–1.07)	0.219		
Ileum	1.04 (0.85–1.26)	0.725		
Small intestine, not specified	0.86 (0.43–1.73)	0.670		
8th T stage				
T1	Reference		Reference	
T2	0.99 (0.34–2.90)	0.986	1.10 (0.38–3.22)	0.863
T3	2.57 (0.96–6.91)	0.061	2.07 (0.77–5.59)	0.149
T4	4.59 (1.71–12.30)	0.002	3.13 (1.16–8.41)	0.024
8th M stage				
M0	Reference		Reference	
M1	3.72 (3.11–4.44)	<0.001	3.72 (3.11–4.44)	<0.001
Grade				
G1/G2	Reference		Reference	
G3/G4	1.68 (1.44–1.97)	<0.001	1.40 (1.19–1.65)	<0.001
Histology				
Conventional adenocarcinoma	Reference		Reference	
Mucinous adenocarcinoma	1.00 (0.76–1.32)	0.978	1.02 (0.77–1.36)	0.854
Signet ring cell carcinoma	1.98 (1.39–2.81)	<0.001	1.09 (0.75–1.58)	0.638
Examined lymph nodes				
≤16	Reference		Reference	
>16	0.82 (0.69–0.97)	0.022	0.79 (0.66–0.94)	0.009
No. Positive lymph nodes				
0	Reference		Reference	
1–2	1.79 (1.46–2.21)	<0.001	1.40 (1.13–1.73)	0.002
3–8	2.95 (2.42–3.58)	<0.001	2.15 (1.75–2.64)	<0.001
>8	4.21 (3.12–5.69)	<0.001	2.72 (1.97–3.76)	<0.001
Tumor size				
≤5 cm	Reference		Reference	
>5 cm	0.74 (0.62–0.88)	<0.001	0.81 (0.67–0.97)	0.024

The number of examined lymph nodes was not included for the AIC is bigger after included it into nomogram, not the optimal model.

positive association between greater numbers of LNs harvested and advanced stage and larger size of the tumor, which was consistent with results reported by previous studies [6]. We also found younger adult patients to be associated with a greater number of LNs harvested, but the location in the jejunum was associated with fewer LNs harvested. In clinical practice, we could not influence host immunology or tumor biology of SBA, although the better survival of cases with >16 nodes may be explained by favorable tumor biology or more radical surgery. However, we propose that surgeons remove more of the mesentery and that pathologists should look harder for nodes, especially in older patients with tumors located in the jejunum and what appears to be less invasiveness. In those cases, patients could gain better survival and would be less likely to be misdiagnosed with respect to stage.

As a prognostic factor, the number of metastatic lymph nodes was also found to be an important risk factor for survival of carcinomas at many sites, such as the colon, appendix, stomach, esophago-gastric junction, and esophagus [10,16,21]. The number of metastatic lymph nodes has been adopted as nodal (N) classification by the 8th edition of the American Joint Committee on Cancer (AJCC) cancer staging

Table 3
Univariate and multivariate Logistic analysis of clinical characteristics for nodal harvest of small intestine adenocarcinoma.

Characteristics	Univariate analysis		Multivariate analysis	
	HR (95%CI)	P value	HR (95%CI)	P value
Age				
<50	Reference		Reference	
50–75	0.84 (0.63–1.13)	0.247	0.71 (0.52–0.96)	0.029
>75	0.63 (0.43–0.93)	0.019	0.51 (0.34–0.77)	0.001
Gender				
Male	Reference			
Female	1.24 (0.99–1.54)	0.060		
Race				
White	Reference			
Black	0.76 (0.56–1.04)	0.086		
Others	1.27 (0.84–1.93)	0.263		
Location				
Duodenum	Reference		Reference	
Jejunum	0.37 (0.27–0.50)	<0.001	0.37 (0.27–0.51)	<0.001
Ileum	0.96 (0.73–1.26)	0.782	1.01 (0.76–1.35)	0.930
Small intestine, not specified	0.28 (0.08–0.96)	0.042	0.23 (0.07–0.78)	0.019
8th T stage				
T1	Reference		Reference	
T2	4.01 (0.89–18.04)	0.070	3.40 (0.88–18.15)	0.073
T3	3.76 (0.87–16.91)	0.076	4.15 (0.95–18.10)	0.058
T4	5.27 (1.23–22.68)	0.026	5.53 (1.27–24.05)	0.023
8th M stage				
M0	Reference		Reference	
M1	0.54 (0.38–0.77)	0.001	0.60 (0.42–0.88)	0.008
Grade				
G1/G2	Reference		Reference	
G3/G4	1.39 (1.11–1.74)	0.004	1.31 (1.03–1.65)	0.026
Histology				
Conventional adenocarcinoma	Reference			
Mucinous adenocarcinoma	1.11 (0.75–1.65)	0.597		
Signet ring cell carcinoma	1.28 (0.74–2.24)	0.372		
Tumor size				
≤5 cm	Reference		Reference	
>5 cm	1.62 (1.29–2.03)	<0.001	1.69 (1.33–2.16)	<0.001

manual. There was also a significant association between highly positive lymph-node ratio (>50–75%) and decreased survival for SBA patients [22]. The number of assessed lymph nodes was visibly correlated with the 5-year disease-specific survival rates of patients with stage II small bowel adenocarcinomas: >7 lymph nodes with 83%, 1–7 lymph nodes with 69%, and 0 lymph nodes with 44% [20]. We used cross-validation to obtain the best grouping of cases by the number of metastatic LNs. The group with 3 cut-off points was finally selected as optimal for discriminating survival in SBA. The final groups selected were N0 (no LN metastasis), N1 (1–2 metastatic LNs), N2 (3–8 metastatic LNs), and N3 (>8 metastatic LNs). This grouping was also verified to be useful by Kaplan–Meier survival analysis.

In fact, the survival was influenced not only by metastatic LN staging. Tumor grade has been reported to be the single most important prognostic factor in small bowel adenocarcinomas [22]. Meanwhile, other independent risk factors, such as age, tumor size, and M stage, were also significant risk factors for poorer survival [3,14]. After uni- and multi-variate Cox analysis of clinical characteristics for prognosis of SBA, we found that age, T stage, M stage, grade, number of metastatic LNs, and tumor size were associated with prognosis in these 2 cohorts. Larger tumor size was found to be associated with better survival. We analyzed the relationship between size and T, N, and M, and found that >5 cm and <5 cm did not differ with respect to T and N, but the M1 ratio of >5 cm was smaller than that of patients with other tumors, indirectly confirming that >5 cm is gentler than <5 cm. This may be because patients with relatively large tumors undergo more comprehensive resection range and more aggressive late radiotherapy and chemotherapy,

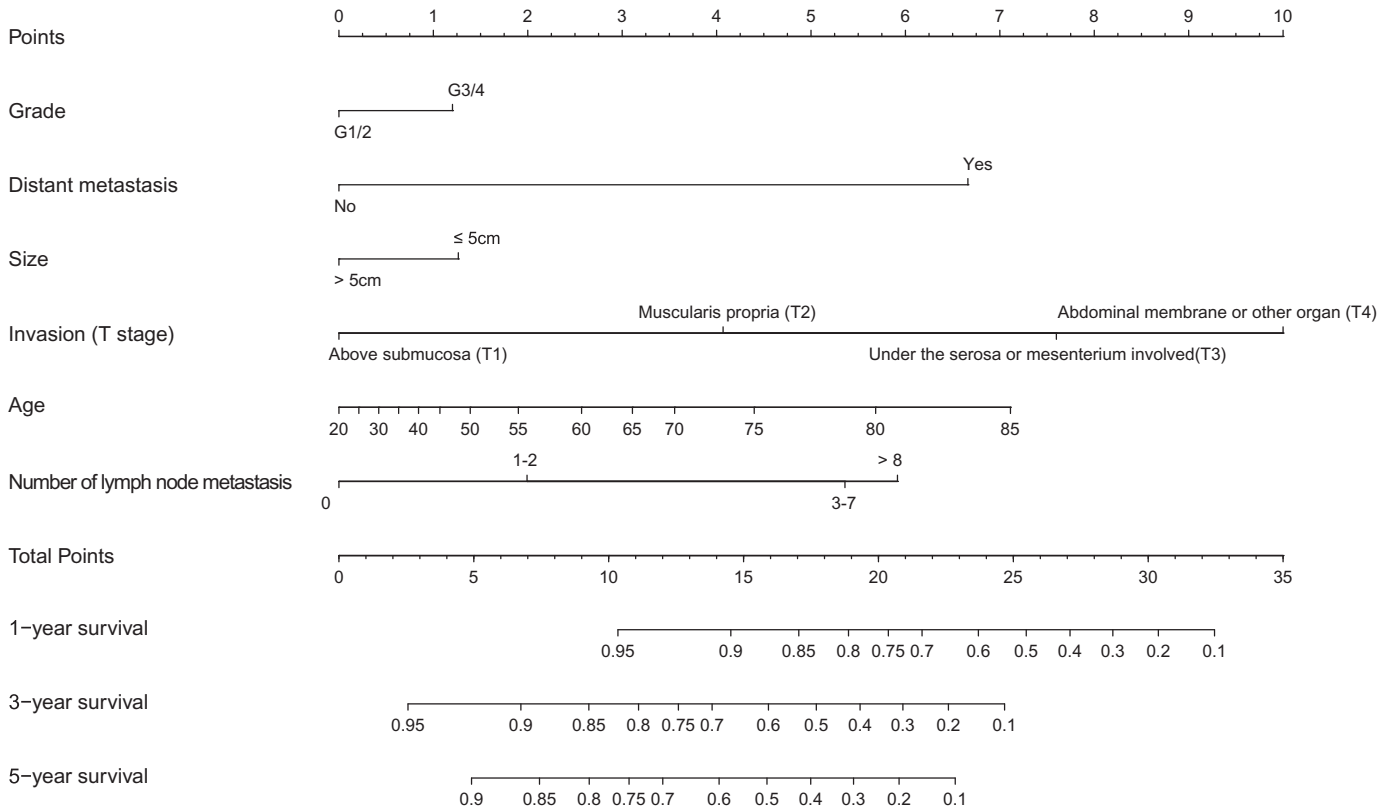


Fig. 3. Nomogram predicted 1- to 5-year cancer specific survival for patients with resected small intestine adenocarcinoma using six available clinical characteristics. For each predictor, read the points assigned on the 0–10 scale at the top and then add these points. Find the number on the “Total Points” scale and then read the corresponding predictions of 1-, 3- and 5-year risk.

which affect prognosis. These may be because tumors in the lumen of the small intestine do not metastasize easily. Primary location was also reported for prognostic value in SBA patients [19]. We analyzed it in the uni- and multi-variate Cox analysis model of clinical

characteristics for the prognosis of SBA and found it to be less important than other factors. We did not include it into the nomogram. In addition, adjuvant chemotherapy could influence the prognosis of SBA patients. It has been reported that adjuvant chemotherapy could significantly

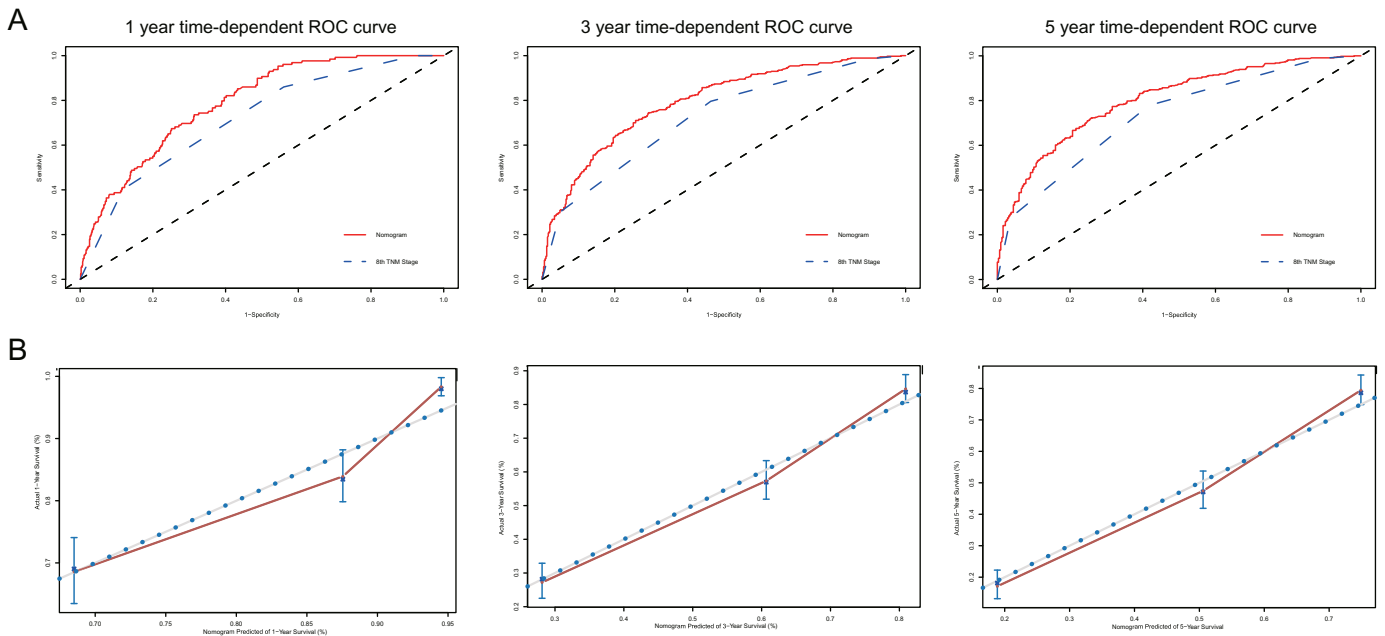


Fig. 4. A. ROC curve of the Nomogram and 8th TNM Stage in prediction of prognosis of patients at 1, 3 and 5 year point in the 2004–2010 cohort. B. The calibration curves for predicting patient survival at 1, 3 and 5 year point in the 2004–2010 cohort. Nomogram-predicted cancer specific survival is plotted on the x-axis; actual cancer specific survival is plotted on the y-axis. A plot along the 45-degree line would indicate a perfect calibration model in which the predicted probabilities are identical to the actual outcomes. ROC: receiver operating characteristic curve; AUC: areas under the ROC curve.

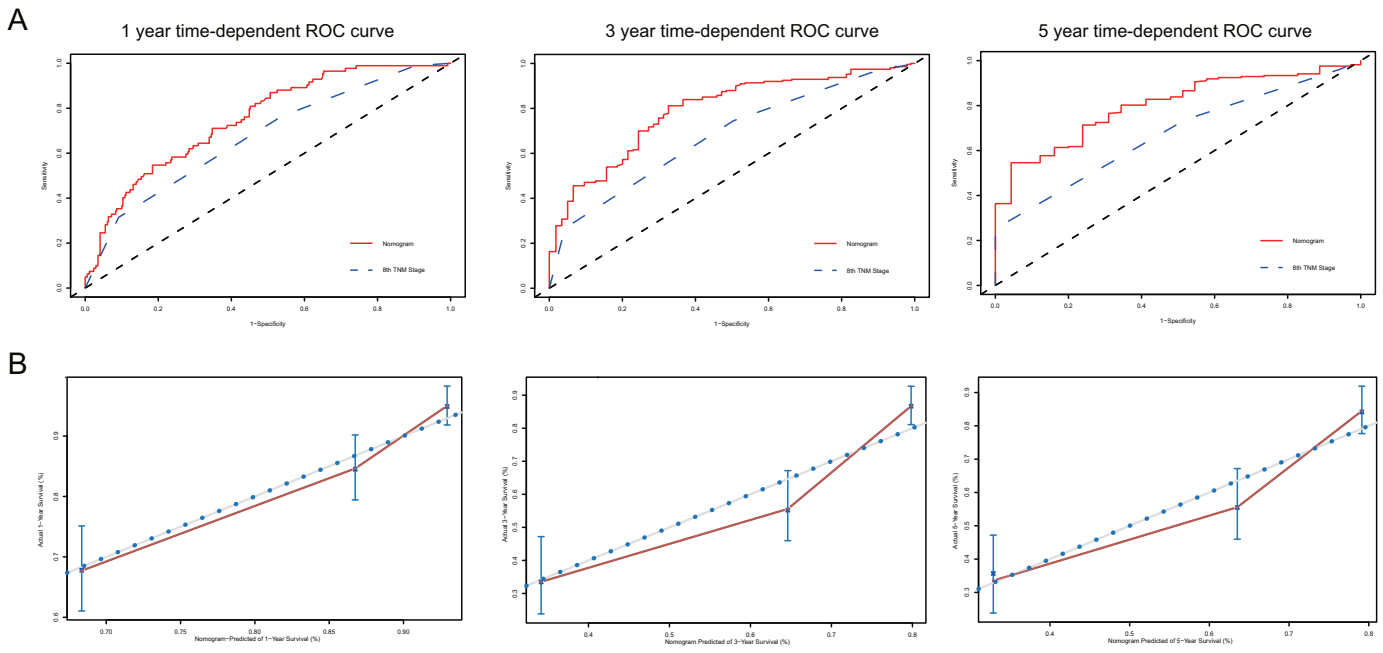


Fig. 5. A.ROC curve of the Nomogram and 8th TNM Stage in prediction of prognosis of patients at 1, 3 and 5 year point in the 2011–2014 cohort. B. The calibration curves for predicting patient survival at 1, 3 and 5 year point in the 2011–2014 cohort. Nomogram-predicted cancer specific survival is plotted on the x-axis; actual cancer specific survival is plotted on the y-axis. A plot along the 45-degree line would indicate a perfect calibration model in which the predicted probabilities are identical to the actual outcomes. ROC: receiver operating characteristic curve; AUC: areas under the ROC curve.

improve stage III SBA patients over with those who received only surgical treatment. However, the efficacy of adjuvant chemotherapy in patients with stage II and poorly differentiated cancer needed to be weighed against potential drug toxicity [7]. Furthermore, biologic prognostic factor also has an effect on the prognosis of SBA. Genomic profiling demonstrated a series of genomic alterations in SBA: APC, CDKN2A, KRAS, SMAD4, BRAF, TP53, PIK3CA, and ERBB2, which may be involved in the development of the disease and the impact of targeted therapy [2,17,24]. The number of lymph nodes examined was not included because the AIC became larger after it was included in the nomogram and found to be sub-optimal. These findings were closely consistent with previous reports on SBA risk factors. Taking the influence of these factors into account, traditional TNM staging might not predict survival well. Combining all of the effective factors, we developed a prediction model using the cohort of patients diagnosed between 2004 and 2010. This model was shown using the nomogram and the validation and calibration of the nomogram was crucial to avoiding model overfitting and determining generalizability. In the current study, the nomogram showed optimal agreement between prediction and actual observation for 1-, 3-, and 5-year survival, which indicated that the

established nomogram was repeatable and reliable. The external cohort of patients diagnosed between 2011 and 2014 for 1-, 3-, and 5-year survival was also well calibrated for the external validation of the nomogram. The validation cohort was not comparable to the internal cohort for the number of lymph nodes examined and the tumor grade. It is difficult to achieve this result under real-world conditions, however, and it is more difficult in the database's retrospective analysis. We grouped patients according to the year in which the tumor was found and analyzed the influencing factors to predict prognosis. By comparing these results to those produced using TNM8, our nomogram prediction model based on the classification of lymph nodes was able to better predict the prognosis in both populations. The model also showed better universality. We also look forward to further validation of this predictive model in forward-looking related research and other large population research.

To evaluate the value of this nomogram for predicting survival, we performed several contrasts with traditional TNM stage. The nomogram produced a higher C-index in both internal cohort and external cohort than TNM stage, which showed greater discriminative ability. The same result was also found in the analysis of specificity using the

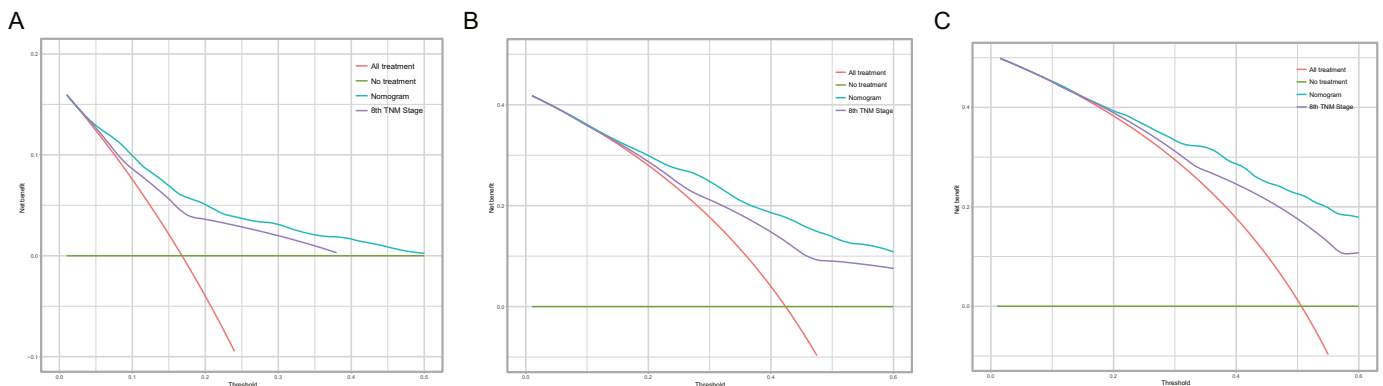


Fig. 6. Decision curve analysis for the Nomogram and the model 8th TNM Stage in prediction of prognosis of patients at 1, 3 and 5 year point.

time-dependent ROC curves for 1-, 3-, and 5-year CSS. It also consistently predicted survival with a higher accuracy than the traditional TNM staging and provided better clinical usefulness throughout the range of survival as assessed by DCA. Through this easy-to-use scoring system, individualized survival prediction could be performed by both physicians and patients after surgery. We found our nomogram to be a more precise prognostic model than the TNM staging system.

The present work has some limitations that should be discussed. These include the retrospective nature of the SEER database collection. The diagnosis of metastatic LN and LN structures all depend on each doctor in different clinical centers. The lack of data for many biological prognostic factors and molecular data could also influence the prognosis of SBA patients. We excluded patients who had missing data regarding any of the collected variables that might have increased the bias. This work is also limited by the failure to incorporate some recognized prognostic parameters or chemotherapy and radiation due to the large bias of the information and lack of detailed information regarding chemotherapy and radiation. Finally, the difference in the number of months of follow-up and the small number of patients followed up for a full 5 years in the validation cohort might have caused bias in predicting the 5-year death risk. Although this nomogram performed well in the two cohorts, it should be used with caution when predicting 5-year risk. We screened the main influencing factors for modeling based on database content. Due to the limitations of the database, the content of the factors named above is not covered and we hope to have relevant data in the future so that we can incorporate it into our research.

In conclusion, this study was performed to examine the number of LNs examined, and we found 17 or more to be optimal. We also modified the current N staging into a 4-level staging system based on the number of metastatic LNs: N0, no LN metastasis; N1, 1–2 metastatic LNs; N2, 3–8 metastatic LNs, and N3, >8 metastatic LNs. Based on this staging, we developed and validated a nomogram with greater clinical usability than traditional TNM staging for the prediction of survival of SBA patients.

Author Contributions

Study concept and design: Z. Ge, X. Li, and S. Wu. Acquisition of data: S. Wu, J. Chen, Q. Zhang and C. Tang. Analysis and interpretation of data: S. Wu, J. Chen, Q. Zhang, M. Tang and X. Zhang. Drafting of the manuscript: S. Wu, Q. Zhang, J. Chen, C. Tang, M. Tang and X. Zhang. Critical revision of the manuscript for important: Z. Ge, X. Li. Statistical analysis: S. Wu, Q. Zhang, J. Chen, C. Tang. Obtained funding: Z. Ge, X. Li.

Funding Sources

This work was supported by grants from the National Natural Science Foundation of China (No.81670505), Shanghai municipal education commission-Gaofeng clinical medicine grant support (DLY201501), Incubating Program for Clinical Research and Innovation of Ren Ji Hospital, School of Medicine, Shanghai Jiao Tong University (Grant No. PYMDT-001), Science and Technology Commission of Shanghai Municipality (Grant No. 16441906903), Three-year action plan for Shin Kang of Shanghai (16CR4005A, 16CR3113B).

Conflict of Interest

There are no conflicts of interest.

Acknowledgements

The authors acknowledge the efforts of the SEER program in the creation of the SEER database.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ebiom.2018.05.022>.

References

- [1] Amin MB, Edge S, Greene F, Byrd DR, Brookland RK. AJCC cancer staging manual. Springer; 2017.
- [2] Aparicio T, Svrcek M, Zaanen A, Beohou E, Laforest A, Afchain P, et al. Small bowel adenocarcinoma phenotyping, a clinicobiological prognostic study. *Br J Cancer* 2013;109:3057–66.
- [3] Bilimoria KY, Bentrem DJ, Wayne JD, Ko CY, Bennett CL, Talamonti MS. Small bowel cancer in the United States: Changes in epidemiology, treatment, and survival over the last 20 years. *Ann Surg* 2009;249:63.
- [4] Blanche P, Dartigues JF, Jacqmingadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat Med* 2013;32:5381.
- [5] Dabaja B, Suki D, Pro B, Bonnen M, Ajani J. Adenocarcinoma of the small bowel: Presentation, prognostic factors, and outcome of 217 patients. *Cancer* 2004;101:518–26.
- [6] Dolan RD, Mccorley ST, Horgan PG, Mcmillan DC. Determinants of lymph node count and positivity in patients undergoing surgery for colon cancer. *Medicine* 2018;97:e0185.
- [7] Ecker BL, McMillan MT, Datta J, Mamtani R, Giantonio BJ, Dempsey DT, et al. Efficacy of adjuvant chemotherapy for small bowel adenocarcinoma: A propensity score-matched analysis. *Cancer* 2016;122:693–701.
- [8] Eo SH, Hong SM, Cho HJ. K-adaptive partitioning for survival data: The kaps add-on package for R; 2013.
- [9] Gerson L, Fidler J, Cave D, Leighton J. ACG clinical guideline: Diagnosis and management of small bowel bleeding. *Am J Gastroenterol* 2015;110:1265–87 [quiz 1288].
- [10] Groth SS, Virnig BA, Whitson BA, Defor TE, Li ZZ, Tuttle TM, et al. Determination of the minimum number of lymph nodes to examine to maximize survival in patients with esophageal carcinoma: Data from the surveillance epidemiology and end results database. *J Thorac Cardiovasc Surg* 2010;139:612–20.
- [11] Haan JC, Buffart TE, Eijk PP, Ma VDW, van Wieringen WN, Howdle PD, et al. Small bowel adenocarcinoma copy number profiles are more closely related to colorectal than to gastric cancers. *Ann Oncol* 2012;23:367.
- [12] Halldanarson TR, McWilliams RR, Donohue JH, Quevedo JF. A single-institution experience with 491 cases of small bowel adenocarcinoma. *Am J Surg* 2010;199:797–803.
- [13] Health NIO. Surveillance, Epidemiology and End Results (SEER) Program; 2011.
- [14] Inoue Y, Hayashi M, Satou N, Miyamoto Y, Hirokawa F, Asakuma M, et al. Prognostic clinicopathological factors after curative resection of small bowel adenocarcinoma. *J Gastrointest Cancer* 2012;43:272–8.
- [15] JA H, BJ M. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 1983;148:839.
- [16] Kang HJ, Eo SH, Kim SC, Park KM, Lee YJ, Lee SK, et al. Increased number of metastatic lymph nodes in adenocarcinoma of the ampulla of Vater as a prognostic factor: A proposal of new nodal classification. *Surgery* 2014;155:74–84.
- [17] Laforest A, Aparicio T, Zaanen A, Silva FP, Didelot A, Desbeaux A, et al. ERBB2 gene as a potential therapeutic target in small bowel adenocarcinoma. *Eur J Cancer* 2014;50:1740–6.
- [18] MB N, V A, WC C, VD V, MS S, G S. Small bowel adenocarcinoma: Understaged and undertreated? *Ann Surg Oncol* 2010;17:2728–32.
- [19] Overman MJ, Hu CY, Kopetz S, Abbruzzese JL, Wolff RA, Chang GJ. A population-based comparison of adenocarcinoma of the large and small intestine: Insights into a rare disease. *Ann Surg Oncol* 2012;19:1439.
- [20] Overman MJ, Hu CY, Wolff RA, Chang GJ. Prognostic value of lymph node evaluation in small bowel adenocarcinoma: Analysis of the surveillance, epidemiology, and end results database. *Cancer* 2010;116:5374–82.
- [21] R V, T S, A K, AB C, AG H. Lymph node evaluation and long-term survival in Stage II and Stage III colon cancer: A national study. *Ann Surg Oncol* 2009;16:585–93.
- [22] Raghav K, Overman MJ. Small bowel adenocarcinomas—Existing evidence and evolving paradigms. *Nat Rev Clin Oncol* 2013;10:534–44.
- [23] Rd HR, Ostrow BH. Adenocarcinoma of the small bowel. *Cancer* 2004;101:518–26.
- [24] Schrock AB, Devoe CE, McWilliams R, Sun J, Aparicio T, Stephens PJ, et al. Genomic profiling of small-bowel adenocarcinoma. *JAMA Oncol* 2017;3:1546–53.
- [25] Venables WN, Ripley BD. Modern applied statistics in S; 2002.
- [26] Young JI, Mongoue-Tchokote S, Wieghard N, Mori M, Vaccaro GM, Sheppard BC, et al. Treatment and survival of small-bowel adenocarcinoma in the United States: A comparison with colon cancer. *Dis Colon Rectum* 2016;59:306–15.