Prognostic Performance of Different Lymph Node Staging Systems After Curative Intent Resection for Gastric Adenocarcinoma

Gaya Spolverato, MD, Aslam Ejaz, MD, MPH,* Yuhree Kim, MD, MPH,* Malcolm H. Squires, MD,† George Poultsides, MD,‡ Ryan C. Fields, MD,§ Mark Bloomston, MD,¶ Sharon M. Weber, MD,|| Konstantinos Votanopoulos, MD,** Alexandra W. Acher, BE,|| Linda X. Jin, BS,§ William G. Hawkins, MD,‡ Carl Schmidt, MD,¶ David A. Kooby, MD,† David Worhunsky, MD,‡ Neil Saunders, MD,** Clifford S. Cho, MD,|| Edward A. Levine, MD,** Shishir K. Maithel, MD,† and Timothy M. Pawlik, MD, MPH, PhD, FACS**

Objective: To compare the prognostic performance of American Joint Committee on Cancer/International Union Against Cancer seventh N stage relative to lymph node ratio (LNR), log odds of metastatic lymph nodes (LODDS), and N score in gastric adenocarcinoma.

Background: Metastatic disease to the regional LN basin is a strong predictor of worse long-term outcome following curative intent resection of gastric adenocarcinoma.

Methods: A total of 804 patients who underwent surgical resection of gastric adenocarcinoma were identified from a multi-institutional database. The relative discriminative abilities of the different LN staging/scoring systems were assessed using the Akaike's Information Criterion (AIC) and the Harrell's concordance index (*c* statistic).

Results: Of the 804 patients, 333 (41.4%) had no lymph node metastasis, whereas 471 (58.6%) had lymph node metastasis. Patients with ≥N1 disease had an increased risk of death (hazards ratio = 2.09, 95% confidence interval: 1.68–2.61; $P < 0.001$]. When assessed using categorical cutoff values, LNR had a somewhat better prognostic performance (C index: 0.630; AIC: 4321.9) than the American Joint Committee on Cancer seventh edition (C index: 0.615; AIC: 4341.9), LODDS (C index: 0.615; AIC: 4323.4), or N score (C index: 0.620; AIC: 4324.6). When LN status was modeled as a continuous variable, the LODDS staging system (C index: 0.636; AIC: 4304.0) outperformed other staging/scoring systems including the N score (C index: 0.632; AIC: 4308.4) and LNR (C index: 0.631; AIC: 4225.8). Among patients with LNR scores of 0 or 1, there was a residual heterogeneity of outcomes that was better stratified and characterized by the LODDS.

Conclusions: When assessed as a categorical variable, LNR was the most powerful manner to stratify patients on the basis of LN status. LODDS was a better predicator of survival when LN status was modeled as a continuous variable, especially among those patients with either very low or high LNR.

From the *Department of Surgery, The Johns Hopkins University School of Medicine, Baltimore, MD; †Department of Surgery, Division of Surgical Oncology, Winship Cancer Institute, Emory University, Atlanta, GA; ‡Department of Surgery, Stanford University, Palo Alto, CA; §Department of Surgery, Washington University School of Medicine, St. Louis, MO; ¶Department of Surgery, The Ohio State University, Columbus, OH; ||Department of Surgery, Division of Surgical Oncology, University of Wisconsin, Madison, WI; and **Department of Surgery, Wake Forest University, Winston-Salem, NC.

Disclosure: A.E. was supported in part by the Eleanor B. Pillsbury Foundation for surgical research. The authors declare no conflicts of interest.

Reprints: Timothy M. Pawlik, MD, MPH, PhD, FACS, Department of Surgery, Johns Hopkins Hospital, 600 North Wolfe Street, Blalock 688, Baltimore, MD 21287. E-mail: tpawlik1@jhmi.edu.

ISSN: 0003-4932/14/26105-0821
DOI: 10.1097/SLA.0000000000001040

Keywords: AJCC staging, gastric adenocarcinoma, LNR, LODDS, lymph node

(*Ann Surg* 2015;262:991–998)

Gastric adenocarcinoma is the fourth most common cancer worldwide and the second most common cause of cancer-related deaths.[1,2] Despite the relatively low incidence of gastric cancer in Western versus Eastern countries, there is an estimated 22,000 new cases and 11,000 deaths related to gastric cancer annually in the United States.[3] The 5-year overall survival for patients with gastric cancer is wide-ranging and estimated to range from 5% to 90%, depending on the extent of disease at presentation.[4–10] Depth of tumor invasion and lymph node (LN) involvement are important prognostic factors associated with overall survival.[11]

In fact, metastatic disease to the regional LN basin is one of the strongest predictors of worse long-term outcome following curative intent resection of most gastrointestinal cancers, including gastric adenocarcinoma.[12–14] Several studies have suggested that survival after surgery is impacted not only by overall LN status (ie, no metastasis vs metastasis) but also the number of LNs with metastatic disease.[15–17] In addition, several studies have suggested that independent of metastatic LN status, a higher number of total LNs examined (TNLE) is associated with survival—a finding that cannot be attributable to understaging.[18,19] In turn, multiple different LN staging/scoring systems have been proposed to stratify patients' long-term prognosis according to LN status.[20–23] Probably, the most widely used staging system is the American Joint Committee on Cancer/International Union Against Cancer (AJCC/UICC) tumor-node-metastasis classification. Only on the basis of the total number of metastatic LNs (NMLNs), the AJCC/UICC staging system does not take into account the number of LNs retrieved at the time of surgery.[24–30] As such, other investigators have proposed that lymph node ratio (LNR), defined as the ratio of NMLNs relative to TNLE, may be a better indicator of the impact of LN status on survival.[31,32] More recently, several investigators have begun to question the accuracy of LNR and have offered alternative LN scoring schemes to predict survival. For example, Wang et al[33] and Sun et al[34] have advocated log odds of metastatic lymph nodes (LODDS), defined as the natural logarithm of the ratio of the probability of an LN to contain metastasis versus the probability of an LN to be free of metastatic disease.[35] In addition, our own group has proposed a novel "N score," which is a prognostic model that takes into account the differential impact of TNLE among patients with and without LN metastasis, as well as the possible nonlinear interaction between TNLE and NMLN.[15]

There is a relative paucity of data regarding the validation of these various LN staging/scoring systems when applied to patients with resected gastric adenocarcinoma.[34,36] Most previous studies have either been small, single institution reports consisting exclusively of Asian patients[37,38] or large reports using administrative data.[36,39] Interpreting the clinical value and use of these prognostic scoring systems among US patients with gastric cancer remains somewhat debatable. Therefore, the objective of this study was to define the relationship of LN status—including NMLN, number of nonmetastatic LN, and TNLE—with survival among patients undergoing surgical resection of adenocarcinoma of the stomach. Specifically, we sought to assess the comparative performance of several of the more widely used LN staging/scoring systems to identify the optimal manner to utilize LN status to predict survival among patients with resected gastric cancer.

METHODS

Patient Selection

Patients who underwent curative intent resection for gastric adenocarcinoma between 2000 and 2012 in 1 of 7 major academic institutions participating in the US Gastric Cancer Collaborative (Johns Hopkins Hospital, Baltimore, MD; Emory University, Atlanta GA; Stanford University, Palo Alto, CA; Washington University, St Louis, MO; Wake Forest University, Winston-Salem, NC; University of Wisconsin, Madison, WI; and The Ohio State University, Columbus, OH) were identified. Patients with gastric cancer other than adenocarcinoma (eg, carcinoid, gastrointestinal stromal tumor, etc) and those who did not undergo an oncological resection with curative intent, or those patients who had missing LN data were excluded from analysis.

Standard data on demographic, clinicopathological, and tumor- and therapy-related variables were collected.[40] Where applicable, information on the type and duration of perioperative chemotherapy and radiotherapy was noted. Data on pathological tumor details, stage, and margin status were obtained. Date of last follow-up and vital status were also collected on all patients. Of particular note, the extent of lymphadenectomy was recorded on the basis of the Japanese Gastric Cancer Association classification system (D1 vs D2 vs D3)[41] and data on TNLE, nonmetastatic LN, and NMLN were also determined. On the basis of this information, data on LN status was categorized for comparative purposes into several different LN staging/scoring systems: seventh edition AJCC/UICC N categories,[42] LNR,[34] LODDS,[39] and N score.[15]

Statistical Analyses

Demographic and clinical characteristics of the study population were reported as numbers (percentage) or medians with interquartile ranges (IQRs). Univariate and multivariate Cox proportional hazard regression models were constructed to explore the association of clinicopathological factors with overall survival. All parameters that were statistically significant in the univariate Cox regression model ($P < 0.05$) were included in the multivariate Cox model. Overall survival was estimated using the Kaplan-Meier method and differences in survival were examined with the log-rank test. In particular, Kaplan-Meier estimates of survival and Cox proportional hazards models were used to explore differences in survival among the strata established by the seventh edition AJCC/UICC N categories,[42] LNR,[34] LODDS,[39] and N score.[15] LNR was evaluated both as a continuous and categorical variable using the cutoff values proposed by Wang et al.[33] Patients with positive LNs were classified into node ratio group 1 (Nr1) to node ratio group 4 (Nr4) on the basis of the following intervals: Nr1: $0 < \text{LNR} \leq 1/15$; Nr2: $1/15 < \text{LNR} \leq 3/10$; Nr3: $3/10 < \text{LNR} \leq 7/10$; and Nr4: $\text{LNR} > 7/10$.

LODDS was calculated using an established formula and was similarly analyzed as both a continuous and a categorical variable using the cutoff values advocated by Sun et al.[34] The N score was derived using the formula proposed by Gleisner et al[15] and assessed as both a continuous and a categorical variable. For the purposes of analyses, N score cutoff values were partitioned into several categories: >2, 2 to 4, 4 to 8, 8 to 12, and >12. Finally, the impact of TNLE on prognosis was determined using different TNLE cutoff points: 0 to 10, 11 to 15, more than 15 LNs.[34]

The relative discriminative abilities of the different LN staging/scoring systems were assessed using the Akaike's Information Criterion (AIC) and the Harrell's concordance index (*c* statistic), a generalization of the area under the receiver operating characteristic curve that quantifies the proportion of all patient pairs for whom the predicted and observed survival outcomes are concordant.[43] A value of $c = 0.5$ indicates no predictive ability as compared with chance alone and a value of 1 indicates perfect discrimination. In general, a predictive model with a low AIC indicates a better model fit and a high *c* statistic represents a better discrimination ability. All analyses were carried out with STATA version 12.0 (StataCorp, LP, College Station, TX). All tests were 2-sided and a $P > 0.05$ was considered statistically significant.

RESULTS

Patient Characteristics and Impact on Overall Survival

Overall, 804 patients with gastric adenocarcinoma who underwent curative intent resection and who met inclusion criteria were identified. Median age was 66.1 years (IQR: 56.7–74.3) and the majority of patients were male (n = 468, 58.2%). Approximately 1 in 5 patients (n = 164, 20.4%) received neoadjuvant chemotherapy before resection, with more patients receiving neoadjuvant therapy in the recent time period (2000–2006, 6.3% vs 2007–2013, 29.4%; $P < 0.001$). Most patients had tumors in the gastric antrum or body (n = 293, 37.2% and n = 283, 35.9%) with a small portion of patients having tumors at the gastroesophageal junction (n = 57, 7.2%). At the time of surgery, patients underwent either a partial (n = 468, 58.5%) or total (n = 332, 41.5%) gastrectomy. On pathology, most tumors were solitary (n = 782, 97.2%) and the median tumor size was 4.0 cm (IQR: 2.5–6.5). Tumors were most often diffuse or mixed-type (n = 370, 68.3%) with the remaining being of the intestinal type (n = 172, 31.7%).

Of the 804 patients, 333 (41.4%) had no LN metastasis, whereas 471 (58.6%) had LN metastasis. On final pathologic analysis, the TNLE was 17 (IQR: 11–25) (2000–2006, 14 vs 2007–2003, 17; $P < 0.01$). The median TNLE in the node-negative group was 15 versus a median TNLE of 18 in the group with LN metastasis ($P < 0.001$). Among patients with nodal disease, the median number of metastatic LNs was 4 (IQR: 2–9); as such, most patients were either pN1 (n = 153; 32.5%) or pN2 (n = 139; 29.5%), whereas a small subset of patients had pN3a (n = 125; 26.5%) or pN3b (n = 54; 11.5%) disease. Most patients with ≥N1 disease had an LNR less than 0.07 (n = 49; 10.4%); however, some patients had an LNR of greater than 0.07 to 0.3 (n = 177; 37.6%), greater than 0.3 to 0.7 (n = 174; 36.9%) or greater than 0.7 (n = 71; 15.1%). Of note, patients were relatively equally distributed among the different N score categories, whereas they were not equally distributed among the different LODDS groups (Table 1).

With a median follow-up of 21 months, the median and 3- and 5-year survivals were 35.8 months and 49.4% and 38.7%, respectively. Several factors were associated with survival on multivariate analysis. Age was associated with a 2% decrease in survival per

TABLE 1. Type of Nodal Dissection, Number of Positive Nodes, Total Harvested Nodes, and 5-Year Survival According to Different Staging Systems

		Type of Nodal Dissection		NMLN,	TNLE,	5-Year
	N (%)	D1 (%)	D2 (%)	Median (IQR)	Median (IQR)	Survival, %
AJCC/UICC 7th edition						
N0	333 (41.4)	125 (67.2)	186 (55.9)	0 (0–0)	15 (8–22)	53.3%
N1	153 (19.0)	60 (69.0)	87 (56.9)	1 (1–2)	17 (10–25)	39.6%
N2	139 (17.3)	46 (52.3)	88 (63.3)	4 (3–5)	15 (10–23)	28.8%
N3a	125 (15.6)	36 (40.4)	89 (71.2)	9 (8–11)	19 (14–24)	20.7%
N3b	54 (6.7)	15 (42.9)	35 (64.8)	21 (18–26)	33 (27–42)	14.0%
LN ratio						
0 of >15	161 (20.4)	50 (31.1)	111 (68.9)	0 (0–0)	22 (18–30)	60.1%
0–0.07 or 0 of ≤15	205 (26.0)	81 (39.5)	109 (53.2)	0 (0–0)	11 (6–15)	45.5%
0.07–0.3	177 (22.5)	54 (30.5)	119 (67.2)	3 (2–4)	18 (13–26)	42.4%
0.3–0.7	174 (22.1)	61 (35.1)	108 (62.1)	8 (5–11)	17 (11–25)	21.9%
>0.7	71 (9.0)	30 (42.3)	36 (50.7)	13 (7–20)	14 (9–22)	4.9%
N score						
<2	186 (23.2)	54 (41.2)	131 (70.4)	0 (0–0)	23 (19–32)	58.9%
2–4	187 (23.3)	78 (72.9)	107 (57.2)	0 (0–2)	13 (10–20)	48.0%
4–8	186 (23.2)	66 (64.7)	102 (54.8)	2 (1–4)	11 (5–17)	35.3%
8–12	113 (14.1)	44 (75.9)	58 (51.3)	6 (3–8)	14 (6–20)	31.8%
≥12	130 (16.2)	39 (45.3)	86 (66.2)	13 (10–20)	21 (15–29)	10.8%
LODDS						
≤−1.5	459 (57.2)	163 (58.4)	279 (60.8)	0 (0–1)	17 (11–25)	49.9%
−1.5 to −1.0	61 (7.6)	17 (39.5)	43 (70.5)	4 (3–5)	18 (13–26)	46.5%
−1.0 to −0.5	63 (7.9)	21 (52.5)	40 (63.5)	5 (3–8)	14 (10–25)	28.8%
−0.5 to 0	87 (10.8)	29 (60.4)	48 (55.2)	7 (3–10)	15 (6–24)	28.0%
>0	132 (16.5)	51 (68.9)	74 (56.1)	11 (7–18)	16 (10–24)	10.3%

patient year [hazards ratio (HR) = 1.02, 95% confidence interval (CI): 1.01–1.03; $P < 0.001$]. Patients with a T3 tumor had a 2-fold increase in the risk of death (HR = 2.08, 95% CI: 1.40–3.09), whereas patients with T4 tumors had a near threefold increased risk (HR = 2.98, 95% CI: 1.96–4.54) compared with patients who had T1 tumors (both $P < 0.001$). Gastroesophageal tumor location was also associated with worse survival, even after adjusting for T stage (HR = 1.68, 95% CI: 1.03–2.73; $P = 0.036$).

Impact of LN Status on Risk of Death: Performance of Various LN Staging/Scoring Systems

LN status was strongly associated with prognosis. Specifically, patients with N0 disease had a median survival of 68.2 months compared with a median survival of 23.8 months for patients with ≥N1 disease (HR = 2.07, 95% CI: 1.66–2.59; $P \le 0.001$). Both TNLE and NMLN were associated with long-term outcome. Specifically, among patients with no LN metastasis, the risk of death increased considerably for each LN examined up to 20 TNLE (HR = 1.12, 95% CI: 1.08–1.15; $P < 0.001$). Among patients with LN metastasis, the risk of death increased with the NMLN, independent of TNLE (HR = 1.07, 95% CI: 1.05–2.00; $P < 0.001$). When the TNLE was low (ie, <10), however, the risk of death was higher than predicted by the independent effects of either TNLE or NLMN.

Survival data according to the 4 LN staging/scoring systems are presented in Figure 1. The seventh edition AJCC/UICC N system was able to stratify patients into 5 distinct prognostic cohorts: pN0 53.3%, pN1 39.6%, pN2 28.8%, pN3a 20.7%, and pN3b 14.0% (log rank: $P < 0.001$) (Fig. 1A). The 5-year overall survival of the different LNR categories, according to the classification proposed by Wang et al,[33] was: Nr0 60.1%, Nr1 45.6%, Nr2 42.4%, Nr3 21.9%, and Nr4 4.9% Fig. 1B); of note, the Nr2 and Nr3 classifications failed to prognostically discriminate patients (log rank: $P = 0.27$). The 5-year overall survival of the different LODDS categories was: LODDS1 49.9%, LODDS2 46.5%, LODDS3 28.8%, LODDS4 28.0%, and LODDS5 10.3% (Fig. 1C); patients in the LODDS2 and LODDS3 categories had overlapping survival curves (log rank: $P = 0.46$). Finally, the 5-year overall survival of the different N score stages based on the classification by Gleisner et al.[15] was: N score 0–2 58.9%, N score 2–4 48.0%, N score 4–8 35.3%, N score 8–12 31.7%, N score >12 10.8% (Figure 1D); using the N score, patients with N score 4 to 8 and 8 to 12 had similar survival curves (log rank: $P = 0.19$) (Table 2).

Through regression modeling, the LN staging/scoring system with the best prognostic discriminatory ability was then assessed through iterative statistical models and comparison of AIC and *c*-statistic values. When assessed using the established categorical cutoff values, LNR was noted to have a somewhat better prognostic performance (C index: 0.630; AIC: 4321.9) than the N score (C index: 0.620; AIC: 4324.6); the AJCC seventh edition (C index: 0.615; AIC: 4341.9) and LODDS (C index: 0.615; AIC: 4323.4) were comparable, both performing slightly less well than LNR and the N score. When stratified by TNLE, LNR remained the best performing model among patients with less than 10 LNs examined (C index: 0.592; AIC: 916.9) and more than 15 LNs examined (C index: 0.641; AIC: 1906.2); interestingly, among patients with 11 to 15 nodes examined, LODDS was the best prognostic model (C index: 0.642; AIC: 672.2).

To assess whether the relative performance of the different LN staging/scoring systems was impacted by the chosen categorical cutoff values, repeat analyses were performed using continuous variables in the statistical models. When LN status was modeled as a continuous variable, the LODDS staging system (C index: 0.636; AIC: 4304.0) outperformed other staging/scoring systems including the N score (C index: 0.632; AIC: 4308.4) and LNR (C index: 0.631; AIC: 4225.8) (Table 3). Of note, all 3 LN scoring schemas performed superiorly to the AJCC seventh edition N staging system (C index: 0.625; AIC: 4341.9). Although LODDS had the highest discriminatory power overall, among patients who had less than 10 TNLE,

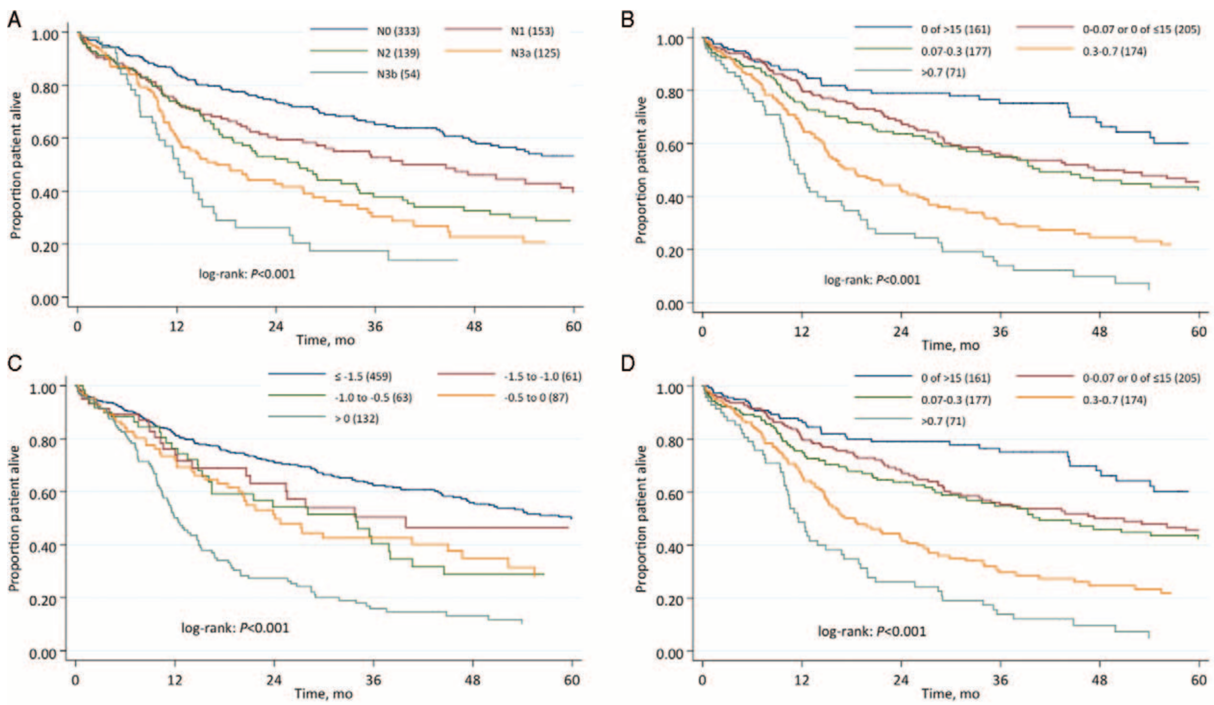

FIGURE 1. Kaplan-Meier curves for overall survival stratified by LN categories based on the AJCC/UICC (A) seventh edition, LNR[33] (B), LODDS (C), and N score (D).

LNR performed the best (LODDS, C index: 0.593; AIC: 987.7 vs LNR C index: 0.594; AIC: 915.3).

Scatter plots were created to evaluate the relationship between LODDS, LNR, and the N score staging systems (Fig. 2). Specifically, to understand the relationship of LODDS to the AJCC/UICC and LNR classifications, scatter plots of the relationship between LODDS and the number or the ratio of LN metastasis were created, as previously described by Sun et al.[34] The value of LODDS

TABLE 2. Overall Survival Rates on the Basis of N Stage, LN Ratio, and N Score Classification According to the LODDS Staging System

	LODDS 1		LODDS 2		LODDS 3		LODDS 4		LODDS 5		
	No.	5-yrs (%)	No.	5-yrs (%)	No.	5-yrs (%)	No.	5-yrs (%)	No.	5-yrs (%)	*P*
N stage											
N0	313	53.2	4	—	—	—	14	58.0		—	0.058
N1	119	43.7	11	35.1	12	24.7	6	—	5	—	0.197
N2	27	36.3	38	51.1	32	30.0	18	8.3	24	13.9	0.068
N3a	—	—	8	41.7	18	39.6	39	29.2	60	9.3	0.023
N3b	—	—	—	—	1	—	10	3.9	43	7.5	0.073
P	0.104		0.221		0.958		0.493	—	0.314		
LN ratio											
0 of >15	161	60.1	—	—	—	—	—	—			NA
0–0.07 or 0 of ≤15	201	46.2	4	—		—	—	—			0.057
0.07–0.3	97	42.5	57	49.0	23	32.3	32.3	—			0.738
0.3–0.7	—	—	—	—	40	26.1	73	21.8	61	18.9	0.076
>0.7	—	—							71	4.9	NA
P	0.018		0.171		0.576	NA		0.342			
N score											
<2	186	58.9	—	—	—	—	—	—			NA
2–4	187	48.0	—	—	—	—	—	—			NA
4–8	86	39.2	57	45.5	34	28.8	9	—			0.097
8–12	—	—	4	50.0	29	28.2	53	38.1	27	18.2	0.268
≥12	—	—					25	32.1	105	70.4	0.056
P	0.056		0.377		0.963		0.233		0.106		

yrs indicates years.

TABLE 3. Prognostic Performance of Different Lymph Node Staging Systems Before and After Stratifying for Total Number of Lymph Nodes Harvested

	No. LNs Recovered							
	All		≤10		11–15		>15	
	C Index	AIC	C Index	AIC	C Index	AIC	C Index	AIC
AJCC/UICC 7th edition	0.615	4341.9	0.582	991.7	0.625	673.5	0.636	1911.4
LN ratio (continuous)	0.631	4225.8	0.594	915.3	0.638	664.7	0.649	1895.9
LODDS (continuous)	0.636	4304.0	0.593	987.7	0.634	663.7	0.649	1895.6
N score (continuous)	0.632	4308.4	0.589	986.6	0.633	665.4	0.644	1900.8
LN ratio (categorical)	0.630	4321.9	0.592	916.9	0.627	672.2	0.641	1906.2
LODDS (categorical)	0.615	4323.4	0.590	992.6	0.642	672.2	0.617	1906.1
N score (categorical)	0.620	4324.6	0.575	992.9	0.623	673.2	0.634	1907.9

increased as the number of LN metastasis increased, suggesting that LODDS and the number of LN metastasis correlated with each other. Interestingly, the correlation was not linear, as the scatter plot relationship plateaued when the TNLE was less than 10. Similarly, the correlation between LODDS and LNR was near, but not perfectly, linear. When LNR was in between 0.2 and 0.8, the LODDS increased much more slowly and plateaued. In contrast, when the LNR was 0 or 1, the relationship was more heterogeneous, suggesting that the LODDS system may have more discriminatory power among patients with either very low or high LNR. In essence, many patients who seemingly had the same prognosis based on either a very low or high LNR, still had a heterogeneous LODDS. These data suggested that the prognosis of patients with LNR of 0 or 1 was not homogeneous and LODDS had more power to discriminate the difference in outcome among patients with LNR of 0 and 1.

DISCUSSION

The accuracy of a staging system to predict long-term survival among patients with cancer is pivotal to help guide postoperative

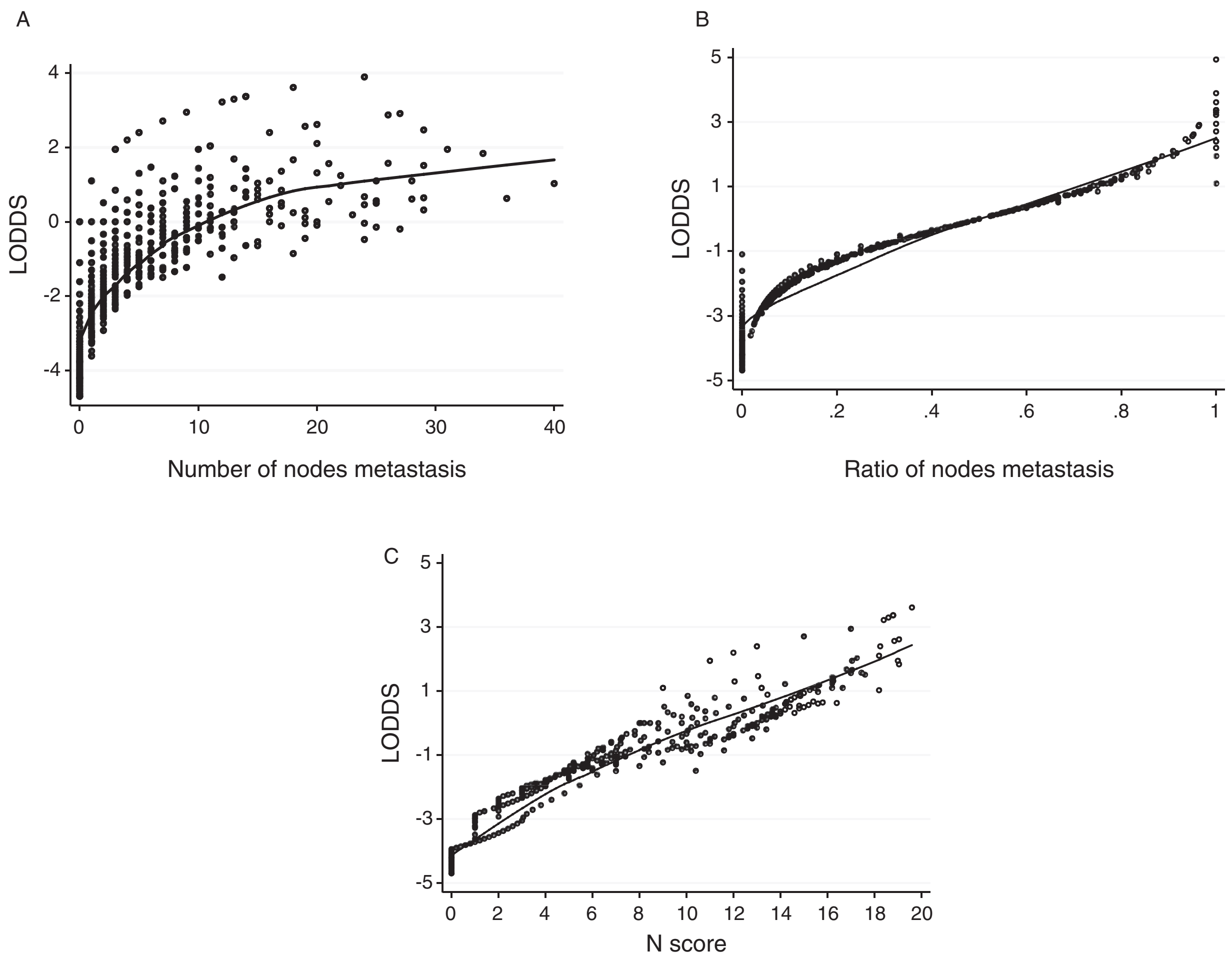

FIGURE 2. Distribution of LODDS vs NMLN (A), LNR (B), and N score (C).

treatment decisions and surveillance. Because of the importance of LN status in determining prognosis after resection of gastric adenocarcinoma, there has been considerable interest in defining an optimal LN staging/scoring system. In fact, over the last decade, a large number of different methods have been proposed to stratify the long-term prognosis of patients after resection of gastric cancer.[32,35,44] Currently, the AJCC seventh edition system is the most widely utilized staging system, but the reliability of the LN staging system has recently come into question.[24–30] Specifically, the AJCC/UICC LN staging system stratifies patients on the basis of the NMLN; however, the NMLN is highly dependent on the TNLE.[39] In turn, both the optimal extent of lymphadenectomy during gastric resection for adenocarcinoma and the average number of LNs harvested at the time of surgery varies widely.[45] As such, there is still significant variability in the extent of LN dissection and average number of nodes retrieved and then analyzed, particularly among patients treated in Western and Eastern centers.[46–49] Given this, multiple alternative staging systems have been proposed to address the shortcomings of the AJCC LN staging system. Several studies have provided mixed results with the evaluations of these different staging systems, with some supporting the prognostic ability of the LNR staging system[36] and other advocating for the use of the LODDS staging system.[34,39] Unlike previous studies, which used administrative data,[36] or were from Eastern institutions,[34] our study is unique in that it was first to evaluate the prognostic ability of 4 of the most commonly cited LN staging systems using a multi-institutional US cohort of patients: AJCC seventh edition, LNR, LODDS, and N score.[15,33,34,44,50] When assessed using the established categorical cutoff values, LNR was noted to have the best prognostic performance, whereas when LN status was modeled as a continuous variable, the LODDS staging system outperformed other staging/scoring systems. Of note, all 3 alternative staging systems (LNR, LODDS, and N score) performed superior to the AJCC seventh edition staging system when analyzed either as categorical or as continuous variables, as well as when stratified by number of LNs retrieved.

The importance of LN involvement in determining prognosis in patients with gastric adenocarcinoma has been well-established.[11] Our data confirm the important prognostic role of LN status while also quantitating the risk of death to increase by 4% with each metastatic LN, even after adjusting for competing risk factors such as age, T stage, and location of the tumor. All 4 staging/scoring systems stratified patients incrementally with regard to long-term outcome (Fig. 1). However, patients who were stage Nr2 versus Nr3 had overlapping survival curves, as did patients with stage LODDS2 and LODDS3 and those patients who were N score 4 to 8 versus 8 to 12 also had overlapping survival curves. These findings suggest that each of the staging systems had some prognostic deficiencies when patients were stratified using specific categorical cutoff values. When data were stratified according to the previously proposed categorical cutoff values, we found that LNR had a somewhat better prognostic performance than LODDS, the N-score, and the AJCC/UICC seventh edition (Table 3). The AJCC/UICC in particular has been criticized for not accounting for the number of LNs actually evaluated. Several authors have noted that a minimal number of LNs need to be evaluated for a patient to be ‘‘adequately’’ staged. For example, Sun et al[34] have advocated that at least 10 LNs be retrieved, whereas AJCC/UICC recommendations consider 15 LNs to be adequate.[50] Some investigators, such as the German Gastric Cancer Study Group, have even suggested a much higher minimum of 25 LNs to achieve accurate staging.[35,51,52] LNR was proposed as a means to take TNLE into account when assessing LN status to avoid the ‘‘Will-Rogers’’ stage migration phenomenon.[31,37,53,54] In this study, we noted that LNR—when examined as a categorical variable—was the best performing model among patients with less than 10 LNs examined. However, the prognostic discriminatory power of LNR seemed poor among patients with either a very low or high LNR (Fig. 2B). These findings may be somewhat intuitive as most clinicians would agree a patient with 1 TNLE and 1 NMLN has a different prognosis compared with a patient with 12 TNLE and 12 NMLN—even though both patients have a LNR of 1.

For patients with either a very low or high LNR, LODDS had a superior discriminatory ability. Although LODDS has been previously examined in a large cohort of patients with colon cancer, the role of LODDS among patients with gastric cancer has not been well studied.[55] In this study, we noted that the heterogeneity of LODDS among patients with LNR scores of 0 or 1 indicated that these patients had varying prognoses despite the fact that they had similar LNR (Fig. 2B). Sun et al[34] demonstrated that LODDS may be a better predictor of survival than LNR. In particular, these investigators found that, whereas the AJCC/UICC and LNR staging systems were dependent on the TNLE, LODDS was able to discriminate patients into distinct prognostic groups, regardless of how many LNs were evaluated. Of note, we also found that LODDS had the highest discriminatory power among patients who had less than 10 TNLE. Collectively, these data strongly suggest that LODDS should be the preferred mechanism to stratify patients with regard to LN status when the TNLE is less than 10. Given that many patients, especially in Western series, have low TNLE, inclusion of LODDS into future editions of the AJCC staging manual should be considered. Although a more complicated calculation than simple number of metastatic nodes or LNR, LODDS seems to be the most reliable method to stratify patients with regard to N classification.

When stratifying patients according to LN status, several different ‘‘ideal’’ cutoffs for each system have been proposed. For example, Zhou et al[56] proposed Nr cutoffs of 0.2 and 0.5, whereas Zeng et al[57] proposed 4 different Nr categories, and Wang et al[39] advocated for 5 categories. The heterogeneity in establishing cutoff values to categorize different LN groups is undoubtedly multifactorial and likely because of the differing statistical power (number of patients), data type (administrative vs clinical data), and the different average number of LN retrieved in these various studies. The lack of consensus on any ideal cutoff value, as well as the fact that NMLN is inherently a discrete integer value (ie, 1, 2, 3, etc), suggest that LN status is better assessed as a continuous variable. Our data support this impression as the prognostic power of all staging systems improved when the LN scores of each system were treated as continuous, rather than categorical, variables. Given this, data from this study emphasize that scoring/staging systems should assess LN status as continuous variables rather than applying indiscriminate categorical cutoffs.

This study has several limitations. As this study was a multi-institutional collaboration among 7 academic institutions across the United States, operative and adjuvant treatment protocols were not standardized. Specifically, with regard to LN retrieval and evaluation, the extent of lymphadenectomy was left to the discretion of the operating surgeon. However, the variable number of TNLE was factored into the analyses and the performance of the different scoring/staging systems were stratified based on TNLE. Finally, although the data from this study were derived from a broad experience of 7 major academic centers across the United States, the findings may not be generalizable to patients treated in different hospital settings.

CONCLUSIONS

Our data suggest that LODDS, N score, and LNR were all more accurate means of predicting long-term survival after curative intent resection of gastric adenocarcinoma versus the AJCC/UICC N staging. When assessed as a categorical variable, LNR was, on average, the most powerful manner to stratify patients on the basis of LN status. However, outcomes among patients with either a very low or high LNR were heterogeneous and poorly characterized by LNR alone. Rather, LODDS was a better predicator of survival. In fact, LODDS was the most accurate scoring system when LN status was modeled as continuous variable. In addition, all staging systems performed better when LN status was treated as continuous variable, indicating that there are likely no ''perfect'' categorical cutoff values for any LN scoring/staging system. Pathological staging of nodal disease is central to prognosis. Data from this study would strongly suggest that the current means of stratifying patients with regard to prognosis and risk of subsequent death using the AJCC method is inadequate. Especially among patients with low TNLE, AJCC—and even LNR—staging may not provide patients and providers with accurate prognostic information. When weighing the risks and benefits of adjuvant therapy, accurate prognostic information is critical to allow for informed treatment decisions. As such, LODDS and N score should be considered when assessing the prognosis of patients with gastric adenocarcinoma to allow a more reliable means to stratify patient survival after surgery.

REFERENCES

1. Jemal A, Bray F, Center MM, et al. *Global cancer statistics*. *CA Cancer J Clin*. 2011;61:69–90.
2. Crew KD, Neugut AI. Epidemiology of gastric cancer. *World J Gastroenterol*. 2006;12:354–362.
3. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. *CA Cancer J Clin*. 2013;63:11–30.
4. Arrington AK, Nelson R, Patel SS, et al. Timing of chemotherapy and survival in patients with resectable gastric adenocarcinoma. *World J Gastrointest Surg*. 2013;5:321–328.
5. Berardi R, Scartozzi M, Romagnoli E, et al. Gastric cancer treatment: a systematic review. *Oncol Rep*. 2004;11:911–916.
6. Degiuli M, Sasako M, Ponti A, et al. Randomized clinical trial comparing survival after D1 or D2 gastrectomy for gastric cancer. *Br J Surg*. 2014;101: 23–31.
7. Dickson JL, Cunningham D. Systemic treatment of gastric cancer. *Eur J Gastroenterol Hepatol*. 2004;16:255–263.
8. Hirota WK, Zuckerman MJ, Adler DG, et al. ASGE guideline: the role of endoscopy in the surveillance of premalignant conditions of the upper GI tract. *Gastrointest Endosc*. 2006;63:570–580.
9. Kim JP, Lee JH, Kim SJ, et al. Clinicopathologic characteristics and prognostic factors in 10 783 patients with gastric cancer. *Gastric Cancer*. 1998;1:125–133.
10. Isozaki H, Okajima K, Fujii K, et al. Effectiveness of paraaortic lymph node dissection for advanced gastric cancer. *Hepatogastroenterology*. 1999;46: 549–554.
11. Posteraro B, Persiani R, Dall'armi V, et al. Prognostic factors and outcomes in Italian patients undergoing curative gastric cancer surgery. *Eur J Surg Oncol*. 2013.
12. Sakuramoto S, Kikuchi S, Futawatari N, et al. Laparoscopy-assisted pancreas- and spleen-preserving total gastrectomy for gastric cancer as compared with open total gastrectomy. *Surg Endosc*. 2009;23:2416–2423.
13. Songun I, Putter H, Kranenbarg EM, et al. Surgical treatment of gastric cancer: 15-year follow-up results of the randomised nationwide Dutch D1D2 trial. *Lancet Oncol*. 2010;11:439–449.
14. de Jong MC, Nathan H, Sotiropoulos GC, et al. Intrahepatic cholangiocarcinoma: an international multi-institutional analysis of prognostic factors and lymph node assessment. *J Clin Oncol*. 2011;29:3140–3145.
15. Gleisner AL, Mogal H, Dodson R, et al. Nodal status, number of lymph nodes examined, and lymph node ratio: what defines prognosis after resection of colon adenocarcinoma? *J Am Coll Surg*. 2013;217:1090–1100.
16. Asiyanbola B, Gleisner A, Herman JM, et al. Determining pattern of recurrence following pancreaticoduodenectomy and adjuvant 5-flurouracil-based chemoradiation therapy: effect of number of metastatic lymph nodes and lymph node ratio. *J Gastrointest Surg*. 2009;13: 752–759.
17. Slidell MB, Chang DC, Cameron JL, et al. Impact of total lymph node count and lymph node ratio on staging and survival after pancreatectomy for pancreatic adenocarcinoma: a large, population-based analysis. *Ann Surg Oncol*. 2008;15:165–174.
18. Parsons HM, Tuttle TM, Kuntz KM, et al. Association between lymph node evaluation for colon cancer and node positivity over the past 20 years. *JAMA*. 2011;306:1089–1097.
19. Baxter NN, Ricciardi R, Simunovic M, et al. An evaluation of the relationship between lymph node number and staging in pT3 colon cancer using population-based data. *Dis Colon Rectum*. 2010;53:65–70.
20. Hermanek P, Sobin LH. *UICC TNM Classification of Malignant Tumors*. 4th ed. Berlin, Germany: Springer; 1987.
21. Japanese Gastric Cancer Association. Japanese classification of gastric carcinoma—2nd English edition. *Gastric Cancer*. 1998; 1:10–24.
22. Greene FL, Page DL, Fleming ID, et al. *AJCC Cancer Staging Manual*. 6th ed. New York, NY: Springer; 2002.
23. Japanese Gastric Cancer A. Japanese classification of gastric carcinoma: 3rd English edition. *Gastric Cancer*. 2011; 14:101–112.
24. Yoo CH, Noh SH, Kim YI, et al. Comparison of prognostic significance of nodal staging between old (4th edition) and new (5th edition) UICC TNM classification for gastric carcinoma. International Union Against Cancer. *World J Surg*. 1999;23:492–497; discussion 497–498.
25. Nio Y, Yamasawa K, Yamaguchi K, et al. Problems in the N-classification of the new 1997 UICC TNM stage classification for gastric cancer: an analysis of over 10 years' outcome of Japanese patients. *Anticancer Res*. 2003;23: 697–705.
26. Shen JY, Kim S, Cheong JH, et al. The impact of total retrieved lymph nodes on staging and survival of patients with pT3 gastric cancer. *Cancer*. 2007;110:745–751.
27. Bouvier AM, Haas O, Piard F, et al. How many nodes must be examined to accurately stage gastric carcinomas? Results from a population based study. *Cancer*. 2002;94:2862–2866.
28. Lee HK, Yang HK, Kim WH, et al. Influence of the number of lymph nodes examined on staging of gastric cancer. *Br J Surg*. 2001;88:1408–1412.
29. Ichikura T, Ogawa T, Chochi K, et al. Minimum number of lymph nodes that should be examined for the International Union Against Cancer(American Joint Committee on Cancer TNM classification of gastric carcinoma. *World J Surg*. 2003;27:330–333.
30. Bertoni E, Zanazzi M, Rosat A, et al. Efficacy and safety of Palmaz stent insertion in the treatment of renal artery stenosis in kidney transplantation. *Transpl Int*. 2000;13 suppl 1:S425-S430.
31. Marchet A, Mocellin S, Ambrosi A, et al. The prognostic value of N-ratio in patients with gastric cancer: validation in a large, multicenter series. *Eur J Surg Oncol*. 2008;34:159–165.
32. Marchet A, Mocellin S, Ambrosi A, et al. The ratio between metastatic and examined lymph nodes (N ratio) is an independent prognostic factor in gastric cancer regardless of the type of lymphadenectomy: results from an Italian multicentric study in 1853 patients. *Ann Surg*. 2007;245:543–552.
33. Wang J, Dang P, Raut CP, et al. Comparison of a lymph node ratio-based staging system with the 7th AJCC system for gastric cancer: analysis of 18,043 patients from the SEER database. *Ann Surg*. 2012;255:478–485.
34. Sun Z, Xu Y, Li de M, et al. Log odds of positive lymph nodes: a novel prognostic indicator superior to the number-based and the ratio-based N category for gastric cancer patients with R0 resection. *Cancer*. 2010;116: 2571–2580.
35. Wang J, Hassett JM, Dayton MT, et al. The prognostic superiority of log odds of positive lymph nodes in stage III colon cancer. *J Gastrointest Surg*. 2008;12:1790–1796.
36. Smith DD, Nelson RA, Schwarz RE. A comparison of five competing lymph node staging schemes in a cohort of resectable gastric cancer patients. *Ann Surg Oncol*. 2014;21:875–882.
37. Liu H, Deng J, Zhang R, et al. The RML of lymph node metastasis was superior to the LODDS for evaluating the prognosis of gastric cancer. *Int J Surg*. 2013;11:419–424.
38. Xu J, Bian YH, Jin X, et al. Prognostic assessment of different metastatic lymph node staging methods for gastric cancer after D2 resection. *World J Gastroenterol*. 2013;19:1975–1983.
39. Wang X, Appleby DH, Zhang X, et al. Comparison of three lymph node staging schemes for predicting outcome in patients with gastric cancer. *Br J Surg*. 2013;100:505–514.
40. Spolverato G, Ejaz A, Kim Y, et al. Rates and patterns of recurrence following curative intent resection for gastric cancer: a united states multi-institutional analysis. *J Am Coll Surg*. 2014;219:664–675.

41. Japanese Gastric Cancer Association. Japanese gastric cancer treatment guidelines 2010 (ver. 3). *Gastric Cancer*. 2011;14:113–123.
42. Edge S, Bryd DR, Compton CC, et al. *AJCC Cancer Staging Manual*. 7th ed. New York, NY: Springer; 2010.
43. Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15:361–387.
44. Sun Z, Zhu GL, Lu C, et al. The impact of N-ratio in minimizing stage migration phenomenon in gastric cancer patients with insufficient number or level of lymph node retrieved: results from a Chinese mono-institutional study in 2159 patients. *Ann Oncol*. 2009;20:897–905.
45. Allum WH. Optimal surgery for gastric cancer: is more always better? *Recent Results Cancer Res*. 2012;196:215–227.
46. Siewert JR, Bottcher K, Roder JD, et al. Prognostic relevance of systematic lymph node dissection in gastric carcinoma. German Gastric Carcinoma Study Group. *Br J Surg*. 1993;80:1015–1018.
47. Wu CW, Hsiung CA, Lo SS, et al. Nodal dissection for patients with gastric cancer: a randomised controlled trial. *Lancet Oncol*. 2006;7:309–315.
48. Cuschieri A, Weeden S, Fielding J, et al. Patient survival after D1 and D2 resections for gastric cancer: long-term results of the MRC randomized surgical trial. Surgical Co-operative Group. *Br J Cancer*. 1999;79:1522–1530.
49. Bonenkamp JJ, Hermans J, Sasako M, et al. Extended lymph-node dissection for gastric cancer. *N Engl J Med*. 1999;340:908–914.
50. Jarnagin WR, Fong Y, Ky A, et al. Liver resection for metastatic colorectal cancer: assessing the risk of occult irresectable disease. *J Am Coll Surg*. 1999;188:33–42.
51. Roder JD, Bottcher K, Busch R, et al. Classification of regional lymph node metastasis from gastric carcinoma. German Gastric Cancer Study Group. *Cancer*. 1998;82:621–631.
52. Schwarz RE, Smith DD. Extended lymph node dissection for gastric cancer: who may benefit? Final results of the randomized Dutch gastric cancer group trial. *J Clin Oncol*. 2005;23:5404–5405; author reply 5405.
53. Inoue K, Nakane Y, Iiyama H, et al. The superiority of ratio-based lymph node staging in gastric carcinoma. *Ann Surg Oncol*. 2002;9:27–34.
54. Yoon SS, Yang HK. Lymphadenectomy for gastric adenocarcinoma: should west meet east? *Oncologist*. 2009;14:871–882.
55. Chang YJ, Chang YJ, Chen LJ, et al. Evaluation of lymph nodes in patients with colon cancer undergoing colon resection: a population-based study. *World J Surg*. 2012;36:1906–1914.
56. Zhou Y, Zhang J, Cao S, et al. The evaluation of metastatic lymph node ratio staging system in gastric cancer. *Gastric Cancer*. 2013;16:309–317.
57. Zeng WJ, Hu WQ, Wang LW, et al. Lymph node ratio is a better prognosticator than lymph node status for gastric cancer: a retrospective study of 138 cases. *Oncol Lett*. 2013;6:1693–1700.