Original Article

Determining the optimal number of examined lymph nodes for prognosis in colon cancer: a population-based study stratified by tumor location and T stage

Jiahao Zhou, Tinghan Yang, Xiangbing Deng, Ziqiang Wang

Colorectal Cancer Center, Department of General Surgery, West China Hospital, Sichuan University, Chengdu, China

Contributions: (I) Conception and design: X Deng, Z Wang; (II) Administrative support: Z Wang; (III) Provision of study materials or patients: J Zhou; (IV) Collection and assembly of data: J Zhou; (V) Data analysis and interpretation: J Zhou, T Yang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Prof. Ziqiang Wang, MD, PhD. Colorectal Cancer Center, Department of General Surgery, West China Hospital, Sichuan University, No. 37 Guoxuexiang Street, Chengdu 610041, China. Email: wangziqiang@scu.edu.cn.

Background: Clinical guidelines recommend ≥12 examined lymph nodes (ELNs) for colon cancer staging, but more may be necessary for accuracy. This study utilized nodal staging scores (NSS) to identity the optimal number of ELNs based on tumor location and T stage, and to assess its prognostic impact.

Methods: Data from 80,792 patients in the Surveillance, Epidemiology, and End Results (SEER) database (2004–2014) and 2,300 patients from the West China Hospital (WCH) cohort (2008–2014) with stage I–III resected colon cancer were analyzed. Optimal ELNs was estimated using a β-binomial distribution model, stratified by tumor location (left-sided, LS; right-sided, RS) and T stage. The primary outcome was overall survival (OS). The association between sufficient nodal staging and OS in node-negative patients was validated by multivariate Cox models.

Results: The mean number of ELN was 18.75 in the SEER and 14.58 in the WCH database. There were 57.8% and 48.8% patients who had RS colon cancer in the SEER and WCH database. Fewer T3-4 tumors were observed in the SEER cohort compared to the WCH cohort (68.4% *vs.* 87.2%). Sufficient nodal staging required ≥24 ELNs for T3 tumors, ≥34 nodes for T4 LS tumors, and ≥40 nodes for T4 RS tumors. For T3 lesions, examining 20–29 ELNs were more likely to have node-positive disease [odd ratio (OR) 1.07; 95% confidence interval (CI): 1.01–1.12] compared to patients with 12–15 ELNs. In the T3N0 group, ELN ≥24 was independently associated with better OS in the SEER database [hazard ratio (HR) 0.72; 95% CI: 0.68–0.75], which was validated in the WCH cohort (HR 0.54; 95% CI: 0.38–0.76).

Conclusions: Optimal ELNs for adequate colon cancer staging is related to both T stage and tumor location. We recommend that ≥24 lymph nodes be examined for T3 tumors, ≥34 for LS T4 tumors and ≥40 for RS T4 tumors for sufficient staging.

Keywords: Lymph node evaluation; colon cancer; tumor sidedness; nodal staging score (NSS)

Submitted Jul 29, 2024. Accepted for publication Dec 10, 2024. Published online Jan 06, 2025.
doi: 10.21037/jgo-24-576
View this article at: https://dx.doi.org/10.21037/jgo-24-576

Introduction

Colon cancer is one of the most prevalent cancers and the second leading cause of cancer-related death worldwide (1). The current tumor-node-metastasis (TNM) staging system for colon cancer involves evaluating the invasive depth of the primary tumor, as well as metastasis to regional lymph nodes and distant organs (2). The presence of lymph node metastasis categorizes the disease as stage III and is significantly related to prognosis. Therefore, accurate nodal staging is critical for effective cancer management which directly influences subsequent use of adjuvant chemotherapy.

The median number of examined lymph nodes (ELNs) has significantly increased through extended surgical procedures such as complete mesocolic excision and improved pathologic techniques such as the use of Patent blue V dye staining (3,4). However, some studies have failed to show a significant increase in detecting node-positive colon cancer (stage III disease) (5,6). Despite this, several studies have demonstrated that patients with a higher number of ELNs are associated with an increased survival rate in colon cancer (7-10). For example, a study by Trepanier *et al.*, which analyzed data from 261,423 colon cancer patients from the National Cancer Database, found that patients with ≥24 ELNs had better OS compared to those with <12 or between 12 and 23 ELNs (9). At the present, clinical guidelines recommend a minimum of 12 ELNs to accurately determine the nodal status in colon cancer (11,12).

However, in some situations, 12-node standard may be insufficient in accurately identifying stage N0 colon cancer. For example, a pervious study has developed the nodal staging score (NSS) model to assess adequate staging of node-negative colon cancer and demonstrated that at least 13 nodes for T3 tumor and 21 nodes for T4 tumor are needed to be examined for a 90% NSS (13), a principle that has also been applied to pancreatic cancer, thyroid cancer, and intrahepatic cholangiocarcinoma (14-16). Moreover, tumor location is another potential factor involved in the nodal staging. Since the distribution of lymph nodes in colon varies significantly according to the tumor sidedness (17), the current strategy of examining 12 lymph nodes may not be equally applicable in colon cancers of different primary sites (18). Collectively, regarding the optimized number of ELNs, tumor location and T stage should both be taken into consideration.

This study aimed to estimate the optimal number of ELNs in colon cancer, stratified by tumor location and T stage, based on the risk of false-negative staging. We then investigated the correlation of these findings with prognosis in two independent cohorts of colon cancer patients. We present this article in accordance with the STROBE reporting checklist (available at https://jgo.amegroups.com/article/view/10.21037/jgo-24-576/rc).

Highlight box

Key findings

- Sufficient nodal staging requires ≥24 examined lymph nodes (ELNs) for T3 tumors, ≥34 nodes for left-sided (LS) T4 and ≥40 nodes for right-sided (RS) T4 colon cancer.
- In T3N0 subgroup, ELN ≥24 was independently associated with better overall survival in both the Surveillance, Epidemiology, and End Results and the West China Hospital cohorts.

What is known and what is new?

- Current guidelines recommend examining at least 12 ELNs for colon cancer staging.
- This study provides new evidence that the optimal number of ELNs for accurate staging should be higher, with specific thresholds based on tumor location and T stage.

What is the implication, and what should change now?

- The optimal number of ELNs for colon cancer staging varies by T stage and tumor location.
- The findings suggest that ≥24 lymph nodes should be examined for T3 tumors, ≥34 for LS T4 tumors and ≥40 for RS T4 tumors to improve staging accuracy and patient prognosis.

Methods

Data source and patient cohorts

We obtained data on patients diagnosed with colon cancer from 2004 to 2014 using the Surveillance, Epidemiology, and End Results (SEER) database via SEER*Stat software (version 8.3.5).

Inclusion criteria were: (I) pathologically confirmed stage I–III colon cancer; (II) diagnosed between 2004 and 2014; (III) underwent radical surgery; and (IV) had valid follow-up data. Exclusion criteria included: (I) missing data of lymph nodes; (II) uncertain pathological stage or tumor location; (III) presence of distant metastasis prior to surgery; and (IV) lack of active follow-up or survival status. Information on the number of lymph nodes evaluated, the number of positive lymph nodes, and the site of primary tumor were included for mathematic calculation (hereafter referred to

as the SEER cohort). The NSS model was developed using the SEER database. The follow-up data for patients in the SEER cohort were served as internal validation.

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). With the approval of the Biomedical Ethics Committee of West China Hospital, Sichuan University (2019 No. 194), we systematically reviewed medical records of patients who underwent radical surgeries for colon cancers between 2008 and 2014 at West China Hospital. Informed consent was taken from all the patients. Patients with the same eligibility criteria as described above were included as an external validation cohort (hereafter referred to as the WCH cohort). Clinical features, including age, gender, year of diagnosis, tumor location, tumor differentiation, pathological T and N stage, number of ELNs, and number of positive lymph nodes, were extracted for analysis both from the SEER and WCH cohorts. Follow-up for WCH patients involved regular monitoring of carcinoembryonic antigen (CEA) and carbohydrate antigen 19-9 (CA19-9) levels, as well as chest and abdominopelvic CT scans every six months for the first three years after surgery and annually thereafter. Follow-up was conducted through outpatient visits, phone calls, and mail surveys.

In both cohorts, patients were classified into the right-sided (RS) colon cancer group (caecum, ascending colon, hepatic flexure, and transverse colon) and the left-sided (LS) colon cancer group (splenic flexure, descending colon, and sigmoid colon) based on the origin site of the primary tumor for subsequent analysis. The primary prognostic outcome was overall survival (OS), defined as the interval from the date of surgery to the date of death from any cause or date of last follow-up visit. Secondary outcome was cancer-specific survival, defined as the interval from the date of surgery to the date of death due to colon cancer or metastatic disease.

NSS model development

A previously described mathematical model was adopted to estimate the probability of false-negative staging as a function of the total number of ELNs (13). First, we estimated the probability of false-negative nodal staging, defined as the probability of removing only negative lymph nodes from a node-positive patient (defined as patients with at least one pathologically diagnosed positive nodes). The probability of positive ELNs among node-positive patients with increasing total number of evaluated lymph nodes was fitted using a β-binomial distribution. This method is commonly applied in the analysis of hierarchical binary data and has demonstrated good model fit in previous studies (13-16). Using the R package VGAM, the parameters of the β-distribution with shapes of α and β were estimated through a maximum likelihood approach. These α and β parameters were fitted across all node-positive patients in the SEER cohort. Subsequently, we calculated the probability of missing metastatic nodal disease *P(FNm)* at each possible number of ELN as follows:

$$P(FNm) = P(0|m,\alpha,\beta) = \frac{Beta(\alpha,\beta+m)}{Beta(\alpha,\beta)} \quad [1]$$

Here, *m* represents the number of ELN and *Beta ()* represents the beta distribution function.

Next, we calculated the number of false-negative patients in the SEER cohort. This was calculated as follows:

$$\#FNm = \frac{P(FNm) * \#TPm}{1 - P(FNm)} \quad [2]$$

Here, *#FNm* represents the number of false negative patients when m of ELNs. *#TPm* represents the number of true positive patients when *m* of ELNs.

Then, the corrected prevalence of true node-positive patients was calculated by the number of false negative patients and the number of true positive patients. The corrective prevalence was determined as follows:

$$prev = \frac{\sum_m (\#TPm + \#FNm)}{\sum_m (\#TPm + \#TNm + \#FNm)} \quad [3]$$

where *#TNm* indicated the number of patients with true-negative lymph nodes evaluation where *m* of ELNs.

Finally, we calculated NSS, the probability of a pathologically node-negative patient which is indeed free of lymph node metastasis *P*(*TNm*) by examined number of *m* as follows:

$$P(TNm) = \frac{1 - prev}{1 - prev + [prev * P(FNm)]} \quad [4]$$

Statistical analysis

Categorical data were presented as frequencies and percentages; continuous variables were presented as means with standard deviation (SD) or medians with interquartile

range (IQR). The Chi-squared test, in conjunction with Bonferroni corrections for pairwise analysis, was used to compare the difference of categorical variables across groups. Student's *t*-test was used to compare the means of two groups. The correlation coefficient between two continuous variables was estimated using a linear model. Multivariate logistic regression was used to examine the diagnostic impact of patient clinicopathological characteristics, including lymph node evaluation, on the diagnosis of node-positive disease. Lymph node evaluation was categorized in two ways as previously described (19) with 12 or more ELNs or not, which is defined as adequate lymph node evaluation by current guidelines (model 1); and a series of continuous groups with small intervals (model 2), which enables a more comprehensive evaluation. Moreover, subgroup analyses were performed to examine the associations between node positivity and node evaluations in patients with different T stages instead of in the overall population.

After determining the optimal cutoff for adequate ELN using the NSS model, patients were divided into two groups: those with inadequate ELN (below the cut-off) and those with adequate ELN (equal or above the cutoff). The survival curves between these patients were generated using the Kaplan-Meier method and compared using the log-rank test. A multivariate Cox proportional hazards model was used to calculate the hazard ratio (HR) and 95% confidence interval (CI). We used "VGAM" R package to develop this mathematical model. Statistical analysis was performed using R, version 3.4.1 (R, Foundation for Statistical Computing, Vienna, Austria). A P value <0.05 was considered statistically significant.

Results

Patient characteristics

A total of 80,792 patients from the SEER database and 2,300 patients from the WCH database who meet the eligible criteria were included in this study. The clinicopathological characteristics of patients in the two cohorts were shown in *Table 1*. The mean number of ELNs was 18.75±12.19 and 14.58±10.24 in the SEER cohort and the WCH cohort, respectively. Of these, 31,217 (38.6%) and 884 (38.4%) node-positive cases were found in the SEER cohorts and the WCH cohorts, respectively. In the SEER cohort, the number of ELNs stratified by specific subgroups was detailed in Table S1. Colon cancer located in the hepatic flexure exhibited the highest mean number of ELNs (20.9), whereas those in the sigmoid colon demonstrated the lowest (16.7). Additionally, a significantly higher proportion of patients with ≥24 ELNs were found in RS colon cancers compared to LS (26.8% *vs.* 17.6%).

Association between lymph nodes evaluation and nodal staging

To examine the impact of lymph node evaluation on the diagnosis of node-positive disease in the SEER cohort, two multivariate logistic models using different classifications of lymph node evaluations were adopted (*Table 2*). Overall, those with at least 12 ELNs (adequate lymph node evaluation in model 1) were more likely to be diagnosed with the node-positive disease regardless of the T stage (adjusted OR for ≥12 nodes *vs.* <12 nodes, ranging from 1.10 to 1.36). Compared to patients with 12 to 15 ELNs (adequate lymph node evaluation in model 2), the further increased number of ELNs was not associated with more diagnosis of node-positive disease in T1 and T2 tumors (*Table 2*). In patients with T3 tumors, those with higher number of ELNs were more likely to have node-positive disease compared with those with adequate nodes evaluated (adjusted OR for 20–29 nodes *vs.* 12–15 nodes, 1.07; 95% CI: 1.01–1.12). Although not statistically significant, the same tendency was also observed in patients with T4 tumors (adjusted OR for ≥40 nodes *vs.* 12–15 nodes, 1.15; 95% CI: 0.92–1.45).

Estimation of adequacy of lymph node evaluation

To find the optimal number of ELNs necessary for sufficient nodal staging, we first estimated the probability of false-negative staging as a function of total node number examined using a previously described mathematical model (13). Briefly, two β-binomial distributions with model parameters of α=1.12 (95% CI: 1.15–1.22), 1.38 (95% CI: 1.15–1.22) and β=3.88 (95% CI: 3.76–4.01), 4.37 (95% CI: 4.21–4.54) were developed using the node-positive cases in the SEER cohort (n=31,217) to estimate the probability of false-negative nodal staging in patients with RS and LS, respectively. As shown in *Figure 1A*, for those with known node-positive disease, approximately 20% RS and 18% LS would be misclassified as node-negative disease when 12 ELNs. Next, we stratified these patients by tumor location and T stage. As shown in *Table 3*, the underestimation of node-positive disease was observed in all T subgroups, and the prevalence increased by T stage. In

Table 1 Clinicopathological characteristics of patients

Variable	SEER cohort (n=80,792)	WCH cohort (n=2,300)	P
Male, n (%)	42,246 (52.3)	1,317 (57.3)	<0.001
Age (years), n (%)			<0.001
<50	9,288 (11.5)	525 (22.8)	
50–69	45,637 (56.5)	1,155 (50.2)	
70+	25,867 (32.0)	620 (27.0)	
Year of surgery, n (%)			<0.001
2004	6,924 (8.6)	0 (0.0)	
2005	6,705 (8.3)	0 (0.0)	
2006	7,084 (8.8)	0 (0.0)	
2007	7,548 (9.3)	0 (0.0)	
2008	7,625 (9.4)	20 (0.9)	
2009	7,682 (9.5)	276 (12.0)	
2010	7,243 (9.0)	348 (15.1)	
2011	7,266 (9.0)	379 (16.5)	
2012	7,421 (9.2)	383 (16.7)	
2013	7,275 (9.0)	392 (17.0)	
2014	8,019 (9.9)	502 (21.8)	
T stage, n (%)			<0.001
T1	12,193 (15.1)	71 (3.1)	
T2	13,356 (16.5)	222 (9.7)	
T3	46,267 (57.3)	1,167 (50.7)	
T4	8,976 (11.1)	840 (36.5)	
ELN, mean (SD)	18.75 (12.19)	14.58 (10.24)	<0.001
LNM, mean (SD)	1.46 (3.17)	1.16 (2.47)	<0.001
N stage, n (%)			<0.001
N0	49,575 (61.4)	1,416 (61.6)	
N1	20,235 (25.0)	653 (28.4)	
N2	10,982 (13.6)	231 (10.0)	
Tumor differentiation, n (%)			<0.001
I	7,814 (9.7)	49 (2.1)	
II	58,505 (72.4)	1,541 (67.0)	
III	12,747 (15.8)	710 (30.9)	
IV	1,726 (2.1)	0 (0.0)	
Right-sided, n (%)	46,666 (57.8)	1,122 (48.8)	<0.001

SEER, Surveillance, Epidemiology, and End Results; WCH, West China Hospital; ELN, examined lymph node; LNM, lymph node metastasis; SD, standard deviation.

Table 2 Adjusted odds ratio of node-positive disease in colon cancer patients as determined by multivariate logistic regression analysis in the SEER cohort

Variable	T1		T2		T3		T4	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Nodes examined (*vs.* <12)								
≥12	1.36 (1.20–1.55)		1.22 (1.10–1.36)		1.10 (1.05–1.16)		1.26 (1.12–1.41)	
No. of lymph nodes examined (*vs.* 12–15)								
1–8		0.64 (0.54–0.76)		0.79 (0.68–0.92)		0.82 (0.76–0.88)		0.70 (0.59–0.83)
9–11		0.84 (0.69–1.01)		0.91 (0.78–1.07)		1.04 (0.97–1.12)		0.96 (0.81–1.15)
16–19		0.95 (0.80–1.13)		1.09 (0.96–1.24)		1.02 (0.96–1.08)		1.03 (0.90–1.18)
20–29		0.96 (0.81–1.14)		1.06 (0.94–1.20)		1.07 (1.01–1.12)		1.01 (0.89–1.15)
30–39		0.93 (0.70–1.23)		0.98 (0.81–1.19)		0.98 (0.91–1.05)		1.04 (0.88–1.24)
≥40		0.88 (0.58–1.29)		0.87 (0.66–1.13)		0.94 (0.85–1.03)		1.15 (0.92–1.45)
Age at diagnosis, years (*vs.* <50)								
50–59	0.93 (0.78–1.13)	0.93 (0.78–1.12)	0.79 (0.68–0.92)	0.79 (0.68–0.91)	0.91 (0.85–0.97)	0.90 (0.85–0.96)	0.88 (0.76–1.02)	0.89 (0.77–1.03)
60–69	0.68 (0.56–0.82)	0.68 (0.56–0.82)	0.73 (0.63–0.84)	0.72 (0.63–0.84)	0.77 (0.73–0.82)	0.77 (0.73–0.82)	0.89 (0.77–1.02)	0.90 (0.78–1.03)
70–79	0.62 (0.51–0.76)	0.62 (0.51–0.76)	0.55 (0.48–0.64)	0.55 (0.47–0.64)	0.67 (0.63–0.71)	0.67 (0.63–0.71)	0.73 (0.64–0.84)	0.74 (0.65–0.85)
Gender (*vs.* male)								
Female	0.84 (0.75–0.93)	0.84 (0.75–0.93)	0.98 (0.90–1.06)	0.98 (0.90–1.06)	0.96 (0.93–1.00)	0.96 (0.93–1.00)	0.98 (0.90–1.07)	0.99 (0.90–1.08)
Tumor location (*vs.* right-sided)								
Left-sided	1.43 (1.27–1.61)	1.43 (1.27–1.62)	1.51 (1.38–1.65)	1.51 (1.38–1.66)	1.27 (1.22–1.32)	1.28 (1.23–1.33)		0.99 (0.90–1.08)
Tumor grade (*vs.* grade 1)								
Grade 2	1.78 (1.53–2.10)	1.78 (1.53–2.09)	1.52 (1.30–1.77)	1.52 (1.30–1.77)	1.39 (1.29–1.50)	1.39 (1.29–1.50)	1.28 (1.06–1.55)	1.28 (1.06–1.54)
Grade 3	4.33 (3.52–5.34)	4.33 (3.52–5.34)	3.13 (2.60–3.79)	3.14 (2.60–3.79)	2.65 (2.43–2.89)	2.65 (2.43–2.89)	2.78 (2.28–3.41)	2.76 (2.26–3.38)
Grade 4	2.94 (1.61–5.07)	2.97 (1.63–5.11)	3.59 (2.48–5.16)	3.60 (2.49–5.17)	2.59 (2.24–3.00)	2.59 (2.24–2.99)	2.63 (1.99–3.49)	2.62 (1.98–3.48)
Year of surgery (*vs.* 2004–2006)								
2007–2009	0.96 (0.82–1.12)	0.94 (0.81–1.10)	1.05 (0.93–1.19)	1.04 (0.93–1.18)	1.00 (0.95–1.05)	0.99 (0.94–1.05)	1.00 (0.88–1.13)	0.99 (0.87–1.13)
2010–2012	0.96 (0.82–1.13)	0.95 (0.81–1.11)	1.02 (0.90–1.15)	1.01 (0.90–1.15)	0.94 (0.89–0.99)	0.93 (0.88–0.98)	1.05 (0.93–1.18)	1.04 (0.92–1.18)
2013–2014	0.79 (0.66–0.94)	0.78 (0.65–0.93)	1.01 (0.88–1.15)	1.00 (0.88–1.15)	0.99 (0.93–1.05)	0.98 (0.93–1.04)	1.19 (1.04–1.35)	1.18 (1.03–1.34)

Data are presented as OR (95% CI). Lymph node evaluation was categorized in two ways: with ≥12 examined lymph nodes or not (model 1); a series of continuous groups with small intervals of examined lymph nodes (model 2). SEER, Surveillance, Epidemiology, and End Results; OR, odd ratio; CI, confidence interval.

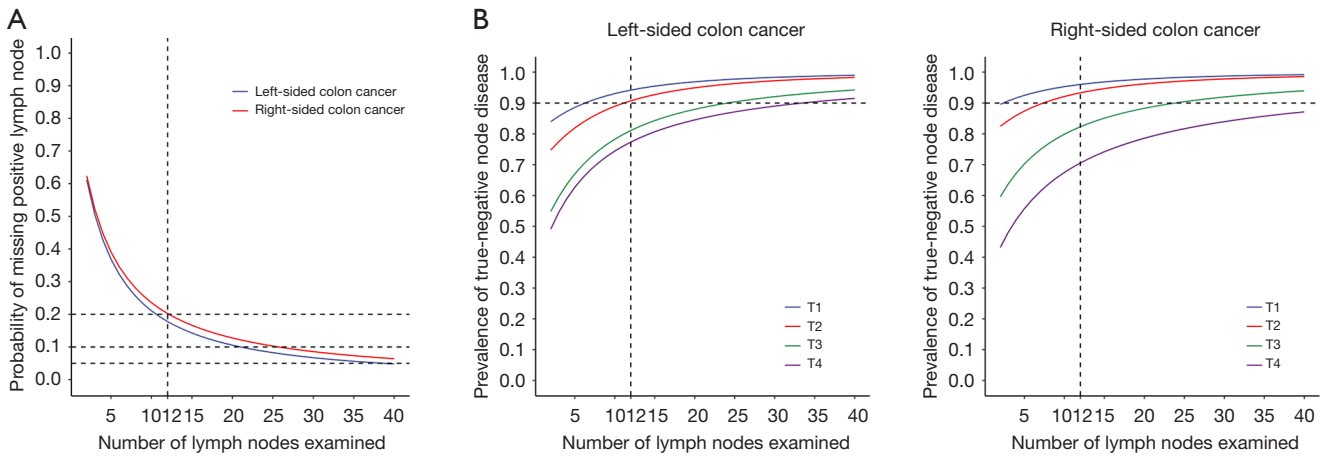

Figure 1 Probability of false-negative node diseases and prevalence of true-negative node diseases by tumor stage as a function of the number of examined lymph nodes. (A) Probability of false-negative node diseases by tumor locations. (B) Prevalence of true-negative node diseases as stratified by tumor locations.

Table 3 Apparent and corrected prevalence of node-positive disease by T stage and tumor locations in the SEER cohort

Group	T stage	Apparent prevalence	Corrected prevalence
Left-sided	T1	0.143	0.203
	T2	0.247	0.317
	T3	0.484	0.570
	T4	0.578	0.659
Right-sided	T1	0.100	0.131
	T2	0.178	0.221
	T3	0.434	0.511
	T4	0.606	0.698

Apparent prevalence was based on the final pathologic stage of the SEER database. Corrected prevalence also considered the probability of false-negative findings based on the number of negative nodes. SEER, Surveillance, Epidemiology, and End Results.

T3 and T4 stage, 8–10% of the patients were estimated to be misdiagnosed with node negative diseases.

Finally, we estimated the probability of true node-negative in the overall cohort when different numbers of lymph nodes were examined (*Figure 1B*). Patients with T1 and T2 tumors, regardless of their tumor locations, will have more than 90% probability of a correct pathological diagnosis of nodal staging with 12 ELNs. However, as also shown in Table S2, using the true node-negative staging with 90% confidence as the threshold, at least 24 lymph nodes were required in T3 patients (LS: 90%, 95% CI: 84.1–93.8%, RS: 90.1%, 95% CI: 84–94%). To achieve the same levels of accuracy, it requires at least 34 ELNs in T4 patients with LS (90.2%, 95% CI: 81.7–95.1%) and ≥40 ELNs for T4 patients with RS (87.1%, 95% CI: 74.2–94.3%).

Risk stratification and validation

The prognostic effect of the optimal number of lymph nodes for sufficient staging was subsequently validated both in the SEER cohort and the WCH cohort. In the multivariate Cox proportional hazards model of OS shown in *Table 4*, the stratification based on at least 24 ELNs remains to be an independent prognostic factor in patients with T3N0 diseases (SEER cohort: HR 0.72, 95% CI: 0.68–0.75, P<0.001; WCH cohort: HR 0.54, 95% CI: 0.38–0.76, P<0.001). Kaplan-Meier curves illustrated the survival benefits from at least 24 ELNs in both cohorts (*Figure 2*). Furthermore, the multivariate Cox proportional hazards model of OS for subgroups consistently demonstrated favorable outcomes for T3N0 patients with ≥24 ELNs in the SEER cohort, regardless of gender, age, and tumor location (Figure S1).

For T4 patients, estimated prevalence of node positive disease was 65.9% and 69.8% in LS and RS (*Table 3*). Namely, patients with T4 tumors should be regarded at ultra-high risk for node-positive disease. As shown in *Figure 3*, survival analysis showed that T4N0 patients had comparable or even worse 5-year outcomes compared to

Table 4 Multivariate Cox proportional hazards model of overall survival for patients with T3N0 diseases

Variable	SEER cohort (n=24,796)	WCH cohort (n=730)
ELNs (*vs*. <24)		
≥24	0.72 (0.68–0.75)	0.54 (0.38–0.76)
Tumor location (*vs*. left-sided)		
Right-sided	1.01 (0.98–1.05)	1.29 (1.03–1.63)
Tumor grade (*vs*. grade 1)		
Grade 2	1.09 (1.02–1.16)	0.55 (0.17–1.72)
Grade 3	1.19 (1.11–1.29)	0.73 (0.23–2.30)
Grade 4	1.19 (1.03–1.38)	N/A
Chemotherapy (*vs*. no)		
Yes	0.64 (0.61–0.67)	N/A

Data are presented as hazard ratio (95% confidence interval). SEER, Surveillance, Epidemiology, and End Results; WCH, West China Hospital; ELN, examined lymph node; N/A, not applicable.

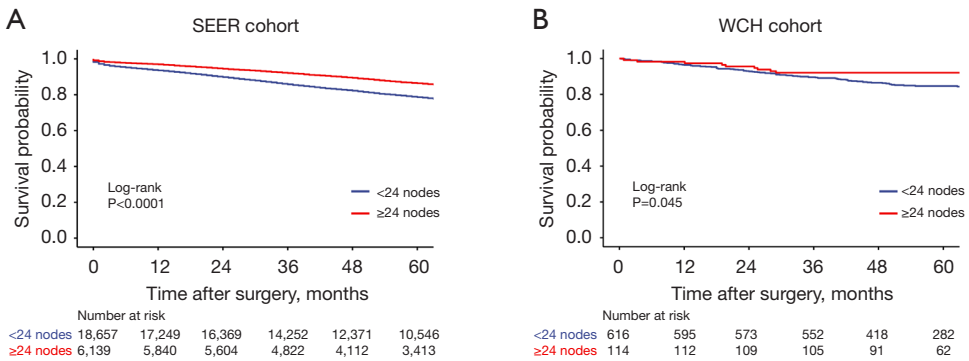

Figure 2 Kaplan-Meier curves of overall survival in T3N0 patients in (A) SEER cohort and (B) WCH cohort. SEER, Surveillance, Epidemiology, and End Results; WCH, West China Hospital.

those with T3N1 tumors. For patients with T4N0 tumors in the SEER cohort (*Figure 3A,3B*), those with sufficient lymph node evaluation (≥34 lymph nodes for LS; ≥40 lymph nodes for RS) were associated with significantly improved OS compared those with insufficient nodal staging. Moreover, as shown in Table S3, better cancer-specific survival outcomes were also observed in the multivariate Cox proportional hazards analysis (LS: HR 0.66, 95% CI: 0.49–0.89; RS: HR 0.75, 95% CI: 0.55–1.03). However, as there were only 1% (2/200) patients with LS tumors and 3.6% (9/246) patients with RS tumors fulfilling the optimized nodes evaluation of T4N0 lesions in the WCH cohort, the survival difference was hard to be calculated for these patients due to the small sample size (*Figure 3C,3D*).

Discussion

In this study, we evaluated the association between lymph node evaluation and nodal staging in patients with colon cancer from 2004 through 2014 using data from the SEER database and the West China Hospital database. While evaluating at least 12 ELNs might lead to more diagnosis of node-positive disease regardless of the T stages, the number of lymph nodes that needs to be evaluated for accurate nodal staging further increases in T3–T4 tumors. We employed a mathematic model based on the number of ELNs to estimate the prevalence of node-positive disease, stratified by tumor locations and T stages. Our results suggested that a cutoff number of 12 lymph nodes is sufficient for nodal

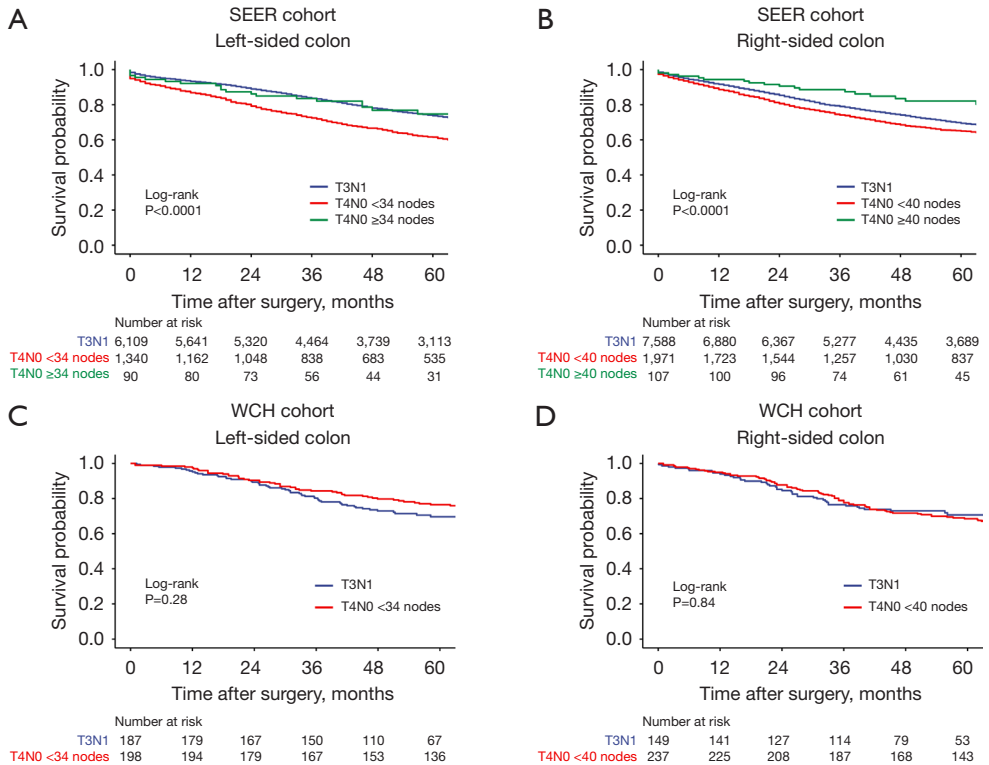

Figure 3 Kaplan-Meier curves of overall survival in T4N0 patients in (A,B) SEER cohort and (C,D) WCH cohort. SEER, Surveillance, Epidemiology, and End Results; WCH, West China Hospital.

staging in T1 and T2 tumors, while a minimum of 24 lymph nodes would be needed to ensure accurate staging in T3 tumors. Moreover, with a high prevalence of node-positive disease, at least 34 and 40 ELNs are needed for accurate staging in T4 disease for LS and RS tumors, respectively. In both the SEER cohort and WCH cohorts, patients with sufficient staging were associated with better OS.

The relationship between lymph node yield and prognosis in colon cancer has been widely discussed (7,20-22), which support the notion that a higher level of lymph node evaluation is associated with improved survival, with the optimal number of ELNs ranging from 16 to 22 nodes. However, most studies have investigated the optimal number of ELNs in the whole population without stratification or only based on single factors (7,20,21). Given the strong correlation between lymph node metastasis and clinicopathological factors (23,24), node evaluation should be optimized not only for the T stage but also for tumor-sidedness. Embryologically, the caecum to the proximal two-thirds of the transverse colon develops from the midgut, distal third of the transverse colon through the upper anal canal develops from the hindgut (25). The molecular features also differ among various locations in colon cancer (26). In addition, RS tumor location is associated with higher lymph node yield (17). The developmental division may lead to divergence in the lymphatic system in each location, implying the necessity for independently optimizing node evaluation.

Postoperative adjuvant chemotherapy is recommended for all node-positive (TNM stage III) patients; however, the effect of adjuvant chemotherapy in stage II colon cancer patients remains controversial (27,28). The definition of high-risk stage II is crucial for postoperative treatment decision-making. Our results, based on the mathematic model, provided a reference for the risk stratification of stage II cases, For T3 lesions, the estimated prevalence of node-positive disease was 57% and 51.1% in LS and RS. The optimal cutoff number of ELNs was 24, based on the mathematic model with 90% confidence of true-negative disease. Subsequent validation also confirms the sensitivity

of this cutoff for risk stratification.

Although the American Joint Committee on Cancer (AJCC) TNM staging system is originally designed to classify prognostic groups, emerging evidence has shown that patients with T4N0M0 (stage IIB–C) tumors have inferior OS compared to those with T1-2N1-2M0 (stage IIIA) tumors (29-31). This disparity has been attributed to the aggressive biology of T4 tumors and the preferential administration of adjuvant chemotherapy for stage IIIA patients due to lymph node metastasis. However, the mechanism underlying this survival paradox remains poorly understood. In this study, we estimated a high prevalence of false-negative staging in patients with T4N0M0 tumors, providing an additional explanation for the survival paradox. Our results suggest that patients with T4 tumors required at least 34 and 40 lymph nodes to achieve accurate nodal staging for LS and RS, respectively. Based on these findings, patients with T4 tumors, particularly those with RS, are at an ultra-high risk for node-positive disease and thus need to be managed appropriately. This may include the adoption of central mesocolic lymph node excision in RS, which yields a significantly higher number of lymph nodes and a lower risk of tumor recurrence, as observed in a population-based cohort study (32), and adjuvant chemotherapy, which has shown exclusively survival benefits in stage II patients with T4 disease rather than those with other high-risk factors (e.g., poor or undifferentiated tumor grade) (28). Further studies on risk stratification of T4N0M0 tumors should also include additional possible risk factors, such as the microsatellite instability (MSI) state and circulating tumor DNA (ctDNA), to facility more informed decision making.

Our study also identified associations between the number of ELNs and patient-level factors, revealing higher ELNs counts among patients with RS tumours, and lower counts among older patients and males. However, non-patient factors may also influence ELN counts. Given the high standardization of surgical techniques such as complete mesocolic excision and Japanese D3 dissection for both RS and LS colon cancer which may have a little impact on lymph node examination (3,33), variations in lymph node examination are more likely attributed to differences in pathological analysis. Indeed, a population-based cohort study highlighted that patients treated by pathologists who were poorer performers in lymph nodes examination were associated with worse prognoses (34).

Several limitations of this study should be noted. Firstly, due to the retrospective nature of this study, potential biases such as selection bias cannot be entirely ruled out, and data on surgical margin status were not available from the SEER database. Secondly, although we reported rates of adjuvant chemotherapy in the SEER cohort, detailed information regarding chemotherapy cycles, regimens and compliance was lacking, leading to uncertainties in the assessment of long-term survival outcomes. Thirdly, despite observing promising outcomes in patients with T4N0 lesions and adequate nodes evaluation, the proportion of this subgroup was low both in the SEER cohort and our hospital, limiting broader clinical application. Additionally, data on tumor microsatellite status were unavailable in the databases. Patients with deficient mismatch repair genes typically exhibit a higher total lymph node yield but a lower likelihood of lymph node metastasis due to strong lymph node reactions (35,36). Higher harvested lymph nodes in such cohorts may also reflect enhanced tumor immune response, potentially favoring immune chemotherapy over standard oxaliplatin-based regimens (37). Finally, our study did not include information on lymph node stations. As Japanese guidelines recommend that metastasis nodes from D3 stations associated with advanced staging (38), analysis based on node locations should be addressed in further studies. Despites these limitations, our study provides valuable insights into the benefits of adequate nodes evaluation for colon cancer patients based on tumor location and stage. Nonetheless, large prospective cohorts or randomized controlled trials are necessary to further validate optimal cutoff values for required lymph nodes.

Conclusions

Although the lymph node distribution varies with tumor location, similar results indicated that similar findings can be achieved for both LS and RS colon cancers. Our analysis suggests that the current cutoff of 12 ELNs is adequate for T1–T2 tumors. However, for T3 tumors, a minimum of 24 ELNs may be necessary to meet standard quality assessment. For patients with T4 tumors, who are at an ultra-high risk for node-positive disease, at least 34 lymph nodes should be examined for LS tumors, and at least 40 lymph nodes should be examined for RS tumors to ensure sufficient staging.

Acknowledgments

None.

Footnote

Reporting Checklist: The authors have completed the STROBE reporting checklist. Available at https://jgo.amegroups.com/article/view/10.21037/jgo-24-576/rc

Data Sharing Statement: Available at https://jgo.amegroups.com/article/view/10.21037/jgo-24-576/dss

Peer Review File: Available at https://jgo.amegroups.com/article/view/10.21037/jgo-24-576/prf

Funding: This study was supported by the National Natural Science Foundation of China (Nos. 82103541, 82203474 and 82203394); Department of Science and Technology of Sichuan Province (Nos. 2022YFS0162, 2021YFS0025, 2022YFS0209 and 2023NSFSC1834); 1·3·5 Project for Disciplines of Excellence, West China Hospital, Sichuan University (Nos. 20HXJS003 and ZYJC21017); 1·3·5 Project for Disciplines of Excellence-Clinical Research Incubation Project, and West China Hospital, Sichuan University (Nos. 22HXFH001 and 2019HXFH031).

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://jgo.amegroups.com/article/view/10.21037/jgo-24-576/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Biomedical Ethics Committee of West China Hospital, Sichuan University (2019 No. 194). Informed consent was taken from all the patients.

References

1. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74:229-63.
2. Gunderson LL, Jessup JM, Sargent DJ, et al. Revised tumor and node categorization for rectal cancer based on surveillance, epidemiology, and end results and rectal pooled analysis outcomes. J Clin Oncol 2010;28:256-63.
3. Hohenberger W, Weber K, Matzel K, et al. Standardized surgery for colonic cancer: complete mesocolic excision and central ligation--technical notes and outcome. Colorectal Dis 2009;11:354-64; discussion 364-5.
4. van Steenbergen LN, van Lijnschoten G, Rutten HJ, et al. Improving lymph node detection in colon cancer in community hospitals and their pathology department in southern Netherlands. Eur J Surg Oncol 2010;36:135-40.
5. Budde CN, Tsikitis VL, Deveney KE, et al. Increasing the number of lymph nodes examined after colectomy does not improve colon cancer staging. J Am Coll Surg 2014;218:1004-11.
6. van Erning FN, Crolla RM, Rutten HJ, et al. No change in lymph node positivity rate despite increased lymph node yield and improved survival in colon cancer. Eur J Cancer 2014;50:3221-9.
7. Men V, Bahl P, Jin JZ, et al. Lymph Node Yield and Long-Term Mortality Risk in Patients with Colon Cancer: A 20-Year Follow-Up National Study. Ann Surg Oncol 2025;32:1117-27.
8. Foo CC, Ku C, Wei R, et al. How does lymph node yield affect survival outcomes of stage I and II colon cancer? World J Surg Oncol 2020;18:22.
9. Trepanier M, Erkan A, Kouyoumdjian A, et al. Examining the relationship between lymph node harvest and survival in patients undergoing colectomy for colon adenocarcinoma. Surgery 2019;166:639-47.
10. Lee L, Erkan A, Alhassan N, et al. Lower survival after right-sided versus left-sided colon cancers: Is an extended lymphadenectomy the answer? Surg Oncol 2018;27:449-55.
11. Benson AB, Venook AP, Adam M, et al. Colon Cancer, Version 3.2024, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw 2024;22:e240029.
12. Weiser MR. AJCC 8th Edition: Colorectal Cancer. Ann Surg Oncol 2018;25:1454-5.
13. Gönen M, Schrag D, Weiser MR. Nodal staging score:

a tool to assess adequate staging of node-negative colon cancer. J Clin Oncol 2009;27:6166-71.
14. Robinson TJ, Thomas S, Dinan MA, et al. How Many Lymph Nodes Are Enough? Assessing the Adequacy of Lymph Node Yield for Papillary Thyroid Cancer. J Clin Oncol 2016;34:3434-9.
15. Hua J, Zhang B, Xu J, et al. Determining the optimal number of examined lymph nodes for accurate staging of pancreatic cancer: An analysis using the nodal staging score model. Eur J Surg Oncol 2019;45:1069-76.
16. Zhang R, Zhang J, Chen C, et al. The optimal number of examined lymph nodes for accurate staging of intrahepatic cholangiocarcinoma: A multi-institutional analysis using the nodal staging score model. Eur J Surg Oncol 2023;49:1429-35.
17. Bilimoria KY, Palis B, Stewart AK, et al. Impact of tumor location on nodal evaluation for colon cancer. Dis Colon Rectum 2008;51:154-61.
18. Yang L, Xiong Z, Xie Q, et al. Prognostic value of total number of lymph nodes retrieved differs between left-sided colon cancer and right-sided colon cancer in stage III patients with colon cancer. BMC Cancer 2018;18:558.
19. 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;306:1089-97.
20. Zhang H, Wang C, Liu Y, et al. The optimal minimum lymph node count for carcinoembryonic antigen elevated colon cancer: a population-based study in the SEER set and External set. BMC Cancer 2023;23:100.
21. Zeng H, Chen Y, Lan Q, et al. Association of hemicolectomy with survival in stage II colorectal cancer: a retrospective cohort study. Updates Surg 2023;75:2211-23.
22. Chen L, Kalady MF, Goldblum J, et al. Does reevaluation of colorectal cancers with inadequate nodal yield lead to stage migration or the identification of metastatic lymph nodes? Dis Colon Rectum 2014;57:432-7.
23. Egger ME, Feygin Y, Kong M, et al. Variation in Lymph Node Assessment for Colon Cancer at the Tumor, Surgeon, and Hospital Level. J Am Coll Surg 2024;238:520-8.
24. Zeng W, Xu J, Liao Z, et al. Construction of a diagnostic nomogram model for predicting the risk of lymph node metastasis in clinical T1 or T2 colon cancer based on the SEER database. Transl Cancer Res 2024;13:1016-25.
25. Cann C, Dotan E. Transverse colon cancer: a call for focused research in an understudied heterogenous disease. J Gastrointest Oncol 2024;15:1981-6.
26. Loree JM, Pereira AAL, Lam M, et al. Classifying Colorectal Cancer by Tumor Location Rather than Sidedness Highlights a Continuum in Mutation Profiles and Consensus Molecular Subtypes. Clin Cancer Res 2018;24:1062-72.
27. Iveson TJ, Sobrero AF, Yoshino T, et al. Duration of Adjuvant Doublet Chemotherapy (3 or 6 months) in Patients With High-Risk Stage II Colorectal Cancer. J Clin Oncol 2021;39:631-41.
28. Verhoeff SR, van Erning FN, Lemmens VE, et al. Adjuvant chemotherapy is not associated with improved survival for all high-risk factors in stage II colon cancer. Int J Cancer 2016;139:187-93.
29. Gunderson LL, Jessup JM, Sargent DJ, et al. Revised TN categorization for colon cancer based on national survival outcomes data. J Clin Oncol 2010;28:264-71.
30. Hari DM, Leung AM, Lee JH, et al. AJCC Cancer Staging Manual 7th edition criteria for colon cancer: do the complex modifications improve prognostic assessment? J Am Coll Surg 2013;217:181-90.
31. Kim MJ, Jeong SY, Choi SJ, et al. Survival paradox between stage IIB/C (T4N0) and stage IIIA (T1-2N1) colon cancer. Ann Surg Oncol 2015;22:505-12.
32. Bertelsen CA, Neuenschwander AU, Jansen JE, et al. 5-year outcome after complete mesocolic excision for right-sided colon cancer: a population-based cohort study. Lancet Oncol 2019;20:1556-65.
33. Banipal GS, Stimec BV, Andersen SN, et al. Are Metastatic Central Lymph Nodes (D3 volume) in right-sided Colon Cancer a Sign of Systemic Disease? A subgroup Analysis of an Ongoing Multicenter Trial. Ann Surg 2024;279:648-56.
34. Becerra AZ, Aquina CT, Berho M, et al. Surgeon-, pathologist-, and hospital-level variation in suboptimal lymph node examination after colectomy: Compartmentalizing quality improvement strategies. Surgery 2017;161:1299-306.
35. Dai X, Dai Z, Fu J, et al. Prognostic significance of negative lymph node count in microsatellite instability-high colorectal cancer. World J Surg Oncol 2024;22:186.
36. Belt EJ, te Velde EA, Krijgsman O, et al. High lymph node yield is related to microsatellite instability in colon cancer. Ann Surg Oncol 2012;19:1222-30.
37. Yu JH, Xiao BY, Li DD, et al. Neoadjuvant camrelizumab plus apatinib for locally advanced microsatellite instability-

high or mismatch repair-deficient colorectal cancer (NEOCAP): a single-arm, open-label, phase 2 study. Lancet Oncol 2024;25:843-52.
38. Noda K, Tominaga T, Nonaka T, et al. Prognostic value of lymph node distribution after laparoscopic colectomy with Japanese D3 dissection. Langenbecks Arch Surg 2024;409:28.

Cite this article as: Zhou J, Yang T, Deng X, Wang Z. Determining the optimal number of examined lymph nodes for prognosis in colon cancer: a population-based study stratified by tumor location and T stage. J Gastrointest Oncol 2025;16(1):115-127. doi: 10.21037/jgo-24-576