



Article Survival Outcome of Gastric Signet Ring Cell Carcinoma Based on the Optimal Number of Examined Lymph Nodes: A Nomogram- and Machine-Learning-Based Approach

Yongkang Lai ^{1,2,†}, Junfeng Xie ^{1,†}, Xiaojing Yin ^{2,†}, Weiguo Lai ¹, Jianhua Tang ¹, Yiqi Du ^{2,*} and Zhaoshen Li ^{2,*}

- ¹ Department of Gastroenterology, Ganzhou People's Hospital Affiliated to Nanchang University, Ganzhou 341000, China
- ² Department of Gastroenterology, Shanghai Changhai Hospital, Naval Medical University, Shanghai 200433, China
- * Correspondence: duyiqi@hotmail.com (Y.D.); zhsl@vip.163.com (Z.L.)
- † These authors contributed equally to this work.

Abstract: The optimal number of examined lymph nodes (ELNs) for gastric signet ring cell carcinoma recommended by National Comprehensive Cancer Network guidelines remains unclear. This study aimed to determine the optimal number of ELNs and investigate its prognostic significance. In this study, we included 1723 patients diagnosed with gastric signet ring cell carcinoma in the Surveillance, Epidemiology, and End Results database. X-tile software was used to calculate the cutoff value of ELNs, and the optimal number of ELNs was found to be 32 for adequate nodal staging. In addition, we performed propensity score matching (PSM) analysis to compare the 1-, 3-, and 5-year survival rates; 1-, 3-, and 5-year survival rates for total examined lymph nodes (ELNs < 32 vs. ELNs \geq 32) were 71.7% vs. 80.1% (p = 0.008), 41.8% vs. 51.2% (p = 0.009), and 27% vs. 30.2% (p = 0.032), respectively. Furthermore, a predictive model based on 32 ELNs was developed and displayed as a nomogram. The model showed good predictive ability performance, and machine learning validated the importance of the optimal number of ELNs in predicting prognosis.

Keywords: gastric signet ring cell carcinoma; examined lymph nodes; minimal number; nomogram; machine learning

1. Introduction

Gastric cancer (GC) is the fifth most frequently diagnosed cancer and the third leading cause of cancer-related deaths worldwide [1]. Gastric signet ring cell carcinoma (GSRC) is an exceptional GC containing a large amount of mucus, accounting for 35–45% of new adenocarcinomas [2]. Compared with non-GSRC, GSRC has a good prognosis in the early stage, while it is characterized by low differentiation, high invasiveness, and poor prognosis in the advanced stage [3,4]. While significant progress has been made in understanding the pathogenesis and molecular biology of GSRC, adequate surgical resection and lymphadenectomy remain the mandatory backbone in treatment with curative intent. However, five-year overall survival after surgery is between 14 and 41.9% [5–8], and distant metastasis may be a primary reason. Therefore, adequate staging information will provide patients with a more accurate prognosis and precise treatment, improving patient survival to a certain extent.

Lymph node staging is crucial in determining the prognosis for GSRC. Several studies have demonstrated the relationship between the number of positive lymph nodes and long-term survival of GSRC patients [5,7,9,10]. However, in clinical practice, the number of positive lymph nodes is closely associated with the number of examined lymph nodes (ELNs), and a lower number of ELNs may result in fewer positive lymph nodes. Inaccurate lymph node staging may frequently occur due to inadequate intraoperative lymph node



Citation: Lai, Y.; Xie, J.; Yin, X.; Lai, W.; Tang, J.; Du, Y.; Li, Z. Survival Outcome of Gastric Signet Ring Cell Carcinoma Based on the Optimal Number of Examined Lymph Nodes: A Nomogram- and Machine-Learning-Based Approach. J. Clin. Med. 2023, 12, 1160. https:// doi.org/10.3390/jcm12031160

Academic Editor: Akira Tsuburaya

Received: 30 December 2022 Revised: 27 January 2023 Accepted: 29 January 2023 Published: 1 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). dissection, affecting the prognosis of patients and treatment decision-making. Therefore, determining the optimal number of ELNs is imperative to improve the prognosis for GSRC.

The optimal number of ELNs for GC has been extensively discussed in the past. According to the National Comprehensive Cancer Network (NCCN) guidelines, GC requires the examination of at least 16 ELNs [11]. According to Hu et al., a minimum of 32 ELNs should be examined in patients with pN3-stage GC [12]. Erstad et al. found that an ELN threshold \geq 30 can significantly improve the prognosis of patients with GC [13]. However, as a more metastatic and more malignant type of GC, GSRC has rarely been studied, and the optimal number of ELNs for GSRC remains controversial. Therefore, the present study aimed to determine the optimal number of ELNs for GSRC; in addition, a prognostic model based on ELNs was developed for predicting the survival outcome for GSRC.

2. Patients and Methods

2.1. Data Source and Study Population

Due to the retrospective nature of the study, informed consent was not required. GSRC cases were obtained from the Surveillance, Epidemiology, and End Results (SEER) Program using SEER*Stat software (version 8.4.0, http://seer.cancer.gov/seerstat/, accessed on 16 May 2022) from the National Cancer Institute. The SEER database collects cancer diagnoses and survival data for about 30% of the US population and benefits from extensive quality review model development. Patients with GSRC diagnosed between 2010 and 2015 who underwent surgery plus lymph node dissection or sampling and had at least one lymph node harvested and examined according to the International Classification of Diseases in Oncology (ICD-O-3, histology code: 8490/3) were included in the study. The total number of regional lymph nodes examined was recorded by the pathologist. The other inclusion criteria were as follows: (1) patients with only one primary tumor; (2) ICD-O-3 identified the primary site as the stomach; and (3) patients with complete survival information.

The exclusion criteria were as follows: (1) unknown number of examined regional nodes; (2) unknown tumor grade, tumor size, or patient race; (3) unknown TNM or American Joint Committee on Cancer (AJCC) stage; (4) unknown surgical information; (5) age < 15 or >95 years old. The variables, including age, sex, race, grade, tumor size, examined lymph node count, tumor location, therapy, AJCC stage, and survival month, were collected from the SEER database. Finally, 1723 patients were included in the study; the detailed screening process for GSRC patients is depicted in Figure 1. Access to the SEER database did not require formal ethical approval, and its open-access policy is included.



Figure 1. Flowchart of patients included in the present study.

2.2. Calculation of Cutoff Value for ELNs

The optimal number of ELNs was determined using X-tile plots. X-tile plots (version 3.6.1, Yale University School of Medicine, New Haven, CT, USA) can divide ELN data into low and high populations and assess all possible divisions [14]. The log-rank survival and means tests were utilized to calculate associations for each division. When a user double-clicked the mouse along the hypotenuse, the software automatically selected the point with the highest chi-square value as the best segmentation of the data. The enumeration method was used to calculate the number of lymph nodes from 15 to 35 for comparison and verify that the chi-square value at this point was the largest. The 1-, 3- and 5-year overall survival rates grouped by the optimal number of ELNs were compared after propensity score matching (PSM) analysis. The propensity model was subjected to PSM analysis in order to control and reduce selection bias and potential confounders. PSM analysis was calculated using a multivariate logistic regression model with the following covariates: age, sex, race, grade, size, site, therapy, and AJCC. The two groups were matched in a 2:1 ratio utilizing the nearest-neighbor method with a caliper width of 0.1. A value of p > 0.05 was considered a baseline balance between the two groups.

2.3. Model Development

Univariate and multivariate Cox regression analyses were used to select the predictive features. A value of p < 0.20 in the univariate analysis was incorporated into the multivariate analysis. The model was developed based on the optimal number of ELNs and features selected in multivariate analysis when p < 0.05. Finally, the optimal number of ELNs, age, treatment, AJCC, tumor size, and tumor site were used to construct a model and presented as a nomogram. To assess the predictive ability of this model in predicting the 1-, 3- and 5-year overall survival rates of patients, receiver operating characteristic curve (ROC) analysis was used, and the area under the ROC (AUC) was calculated. AUC values of 0.5 and 1.0 represented random chance and a significant ability to predict overall survival rates with the model. Calibration curves and decision-curve analysis (DCA) curves were used for the discrimination of the model and to determine the clinical net benefit associated with the use of the model. In addition, Extreme Gradient Boosting (XGBoost) was used to analyze the predictive ability with or without the optimal number of ELNs. XGBoost is a recently developed gradient tree boosting algorithm based on machine learning that combines the outputs of other decision trees to improve classification. XGBoost is scalable and allows faster calculations [15].

2.4. Statistical Analysis

The cutoff value for ELNs was calculated using X-tile software. The other statistical analyses were performed using R statistical software 4.2.0 (www.r-project.org, accessed on 6 July 2022). Continuous variables were expressed as mean \pm standard deviation (SD) for normally distributed data, whereas for non-normally distributed data, continuous variables were expressed as medians (interquartile spacing). Categorical variables were expressed as proportions. The Kaplan–Meier method was used to calculate the survival data, while the log-rank test was further used to assess statistical significance. Variables with *p* < 0.05 were considered statistically significant.

3. Results

3.1. Baseline Characteristics

The data assembly process is depicted in Figure 1. A total of 4652 patients histologically diagnosed with GSRC between 2010 and 2015 were enrolled using the SEER database. Patients with an unknown number of regional nodes examined (n = 369), patients missing tumor grade, race, or tumor size (n = 1498), patients without TNM or AJCC stage information (n = 864), patients without surgical information (n = 148), and patients younger than 18 years old or older than 95 years old (n = 52) were excluded. Ultimately, 1723 patients were included in the study. These patients were randomly partitioned into a training cohort (n = 1206; 70%) and a validation cohort (n = 517; 30%). Detailed information is presented in Table 1. The median age was 63 years (IQR: 53–72), with 50.7% males. By race, white patients (66.8%) accounted for the majority of this study, while black and other patients accounted for 11.8% and 21.4%, respectively. The most common grade was Grade III (94.5%), and 653 (37.9) patients had a tumor size larger than 5 cm. Meanwhile, the gastric antrum was the most common site for GSRC (28.6%). Among all included patients, the 1-, 3-, and 5-year OS rates were 74.2%, 47%, and 29.6%, respectively.

Table 1. Demographic and clinicopathological characteristics of patients.

Variables	All Patients (n = 1723)	Training Cohort (n = 1206)	Validation Cohort (n = 517)
Age (median, IQR)	63 (53–72)	63 (52–72)	62 (52–72)
Sex [No. (%)]			
Male	873 (50.7)	613 (50.8)	260 (50.3)
Female	850 (49.3)	593 (49.2)	257 (49.7)
Race [No. (%)]			
White	1151 (66.8)	797 (66.1)	354 (68.5)
Black	203 (11.8)	145 (12)	58 (11.2)
Other	369 (21.4)	264 (21.9)	105 (20.3)
Grade [No. (%)]			
1	2(0.1)	1 (0.1)	1 (0.2)
	45 (2.6)	32 (2.7)	13 (2.5)
	1628 (94.5)	1141 (94.5)	487 (94.2)
IV	48 (2.8)	32 (2.7)	16 (3.1)
Size [NO. (%)]	117 (6.9)	9E (7)	20 (7 5)
$\leq 1 \text{ cm}$	117(6.8)	00 (7) 172 (14 2)	39 (7.5) 80 (15 E)
1-2 cm	249 (14.3)	175 (14.5)	74(14.2)
2-5 cm	207 (13.3)	160 (13.4)	74(14.3)
4 5 cm	201 (11.6)	100(13.9) 148(12.3)	65 (12.6)
4-5 cm	653 (37.9)	140 (12.3)	183 (35 4)
Examined lymph node count	033 (37.2)	97.1)	105 (55.4)
(vear mean + SD)	20.4 ± 0.33	20.3 ± 0.39	20.7 ± 0.62
Site $[N_0, (\%)]$			
Cardia	239 (13.9)	165 (13.7)	74 (14.3)
Fundus of stomach	42 (2.4)	34 (2.8)	8 (1.6)
Body of stomach	207 (12)	143 (11.9)	64 (12.4)
Gastric antrum	492 (28.6)	342 (28.4)	150 (29)
Pylorus	91 (5.3)	71 (5.9)	20 (3.9)
Lesser curvature of stomach	228 (13.2)	158 (13.1)	70 (13.5)
Greater curvature of stomach	102 (5.9)	73 (6.1)	27 (5.2)
Overlapping	177 (10.3)	115 (9.5)	62 (12)
Stomach NOS	145 (8.4)	103 (8.6)	42 (8.1)
Therapy [No. (%)]			
Surgery only	633 (36.7)	447 (37.1)	186 (36)
Surgery plus radio	22 (1.3)	14 (1.2)	8 (1.5)
Surgery plus chemo	517 (30)	354 (29.3)	163 (31.5)
Surgery plus cadio and chemo	551 (32)	391 (32.4)	160 (31)
AJCC [No. (%)]			
I	379 (22)	273 (22.6)	106 (20.5)
II	374 (21.7)	268 (22.2)	106 (20.5)
III	785 (45.6)	540 (44.8)	245 (47.4)
IV	185 (10.7)	125 (10.4)	60 (11.6)
Overall survival [No. (%)]	1050 (51.0)	000 (51 0)	
1-year	1278 (74.2)	902 (74.8)	3/6 (72.7)
3-year	810 (47)	577 (47.8)	233 (45.1)
5-year	510 (29.6)	355 (29.4)	155 (30)

3.2. Demarcation of the Minimum Number of ELNs

X-tile software determined 32 as the minimum number of ELNs. The Kaplan–Meier survival curves for OS showed that patients with \geq 32 ELNs had better survival times than patients with <32 ELNs (p = 0.032, Figure 2). According to the NCCN guidelines, 16 or more ELNs are recommended. Hence, this study used the enumeration method to calculate the number of lymph nodes from 15 to 35 for comparison (Table 2). When 32 was calculated as the cutoff value for ELNs, the two groups had the maximum chi-square values and relative risk values and minimum *p*-values.



Figure 2. Kaplan–Meier survival curves for overall survival in patients with ELNs < 32 and ELNs \geq 32. ELNs, examined lymph nodes.

Cut off Value for ELNs	Chi-Square Score	Relative Risk	p Value
<15 vs. ≥15	3.02	1.05	0.18
<16 vs. ≥16	2.92	1.04	0.083
<17 vs. ≥17	3.67	1.05	0.088
<18 vs. ≥ 18	3.53	1.06	0.056
<19 vs. ≥19	5.35	1.07	0.061
<20 vs. ≥ 20	4.21	1.06	0.021
<21 vs. ≥ 21	4.45	1.06	0.041
<22 vs. ≥ 22	3.96	1.05	0.035
<23 vs. ≥23	3.53	1.05	0.047
<24 vs. ≥ 24	2.51	1.03	0.58
<25 vs. ≥ 25	4.39	1.05	0.11
<26 vs. ≥ 26	3.79	1.05	0.037
$<27 \text{ vs.} \ge 27$	3.06	1.05	0.052
<28 vs. ≥ 28	3.4	1.06	0.081
<29 vs. ≥29	4.04	1.07	0.066
<30 vs. ≥ 30	3.44	1.07	0.045
<31 vs. ≥31	4.35	1.09	0.063
<32 vs. ≥32	5.45	1.11	0.037
<33 vs. ≥33	3.91	1.10	0.02
<34 vs. ≥ 34	4.69	1.10	0.048
<35 vs. ≥35	3.48	1.10	0.03

Table 2. Analysis of the cutoff value for ELNs.

3.3. Baseline Comparisons on ELNs \geq 32 and ELNs < 32 (Pre-PSM and Post-PSM)

To further verify 32 as the minimum number of ELNs, ELNs \geq 32 and ELNs < 32 were analyzed as baselines for each group (Table 3). Before PSM, there were significant differences in five factors between the two groups, including race (p = 0.015), size (p = 0.002), site (p = 0.003), therapy (p = 0.001), and AJCC (p = 0.001). Moreover, patients with \geq 32 ELNs had better 1-year survival rates than patients with <32 ELNs (80.2% vs. 72.9%, p = 0.01). However, the 3-year and 5-year survival rates between the two groups had no significant difference. After PSM, 548 patients with \geq 32 ELNs were matched with 291 patients with <32 ELNs. There was no significant difference in the baseline characteristics between the two groups, indicating that each group's selection bias was reduced and controlled. As shown in Table 3, patients with \geq 32 ELNs had better 1-year survival rates (80.1% vs. 71.7%, p = 0.008), 3-year survival rates (51.2% vs. 41.8%, p = 0.09), and 5-year survival rates (30.2% vs. 27%, p = 0.032) than patients with <32 ELNs.

Table 3. Clinical characteristics of patients in the study before and after propensity score matching.

Before Matching			After Matching			
Characteristics	ELN < 32 (n = 1430)	$ELN \ge 32$ (n = 293)	p	ELN < 32 (n = 548)	$ELN \ge 32$ (n = 291)	р
Age (median, IQR)	63 (52–73)	60 (50-68)		60 (51–71)	60 (50-68)	0.121
Sex [No. (%)]			0.843			0.640
Male	723 (50.6)	150 (51.2)		288 (52.6)	148 (52.9)	
Female	707 (49.4)	143 (48.8)		260 (47.4)	143 (49.1)	
Race [No. (%)]			0.015			0.519
White	955 (66.8)	196 (66.9)		359 (65.5)	194 (66.7)	
Black	181 (12.7)	22 (7.5)		72 (13.1)	22 (7.5)	
Other	294 (20.5)	75 (25.6)		117 (21.4)	75 (25.8)	
Grade [No. (%)]			0.789			0.988
I	2 (0.1)	0		0	0	
Π	39 (2.7)	6 (2)		9 (1.6)	5 (1.7)	
III	1348 (94.3)	280 (95.6)		525 (95.8)	279 (95.9)	
IV	41 (2.9)	7 (2.4)		14 (2.6)	7 (2.4)	
Size [No. (%)]		()	0.002		()	0.485
<1 cm	107 (7.5)	10 (3.4)		24 (4.4)	10 (3.4)	
$\frac{-}{1-2}$ cm	212 (14.8)	37 (12.6)		53 (9.7)	35 (12)	
2–3 cm	225 (15.7)	42 (14.3)		65 (11.8)	42 (14.4)	
3–4 cm	205 (14.3)	31 (10.6)		77 (14.1)	31 (10.7)	
4–5 cm	169 (11.8)	32 (10.9)		68 (12.4)	32 (11)	
>5 cm	512 (35.9)	141 (48.2)		261 (47.6)	141(48.5)	
Site [No. (%)]	012 (001))	111 (1012)	0.003	_01 (1110)	111 (1010)	0.064
Cardia	210 (14 7)	29 (9 9)	0.000	69 (12 6)	28 (9.6)	0.001
Fundus of stomach	34(24)	8(27)		10 (1.8)	8(27)	
Body of stomach	159(111)	48 (16 4)		60 (10.9)	47(162)	
Castric antrum	427 (29.9)	65 (22 2)		155 (28.3)	65 (22 3)	
Pylorus	81 (57)	10(34)		31 (5 7)	10(34)	
Lesser curvature of stomach	182(12.7)	46 (15 7)		74 (13.6)	46 (15.8)	
Greater curvature of stomach	84 (5 9)	18 (6 1)		27(4.9)	18 (6 2)	
Overlapping	1/1 (0.8)	36 (12 3)		78(14.2)	36(124)	
Stomach NOS	141(9.0) 112(7.8)	33(11.3)		10(14.2)	33(11.4)	
Thorapy [No. (%)]	112 (7.8)	55 (11.5)	0.001	11 (0)	55 (11.4)	0 1 3 0
Surgery only	555 (38.8)	78 (26 6)	0.001	161 (20 4)	76 (26.1)	0.150
Surgery plus radio	18 (1 2)	$\frac{70}{4}(1.4)$		7(12)	4(14)	
Surgery plus facto	10(1.3) 201(27.2)	4(1.4) 126(42)		7 (1.3) 175 (21.0)	4(1.4) 126(42.2)	
Surgery plus chemo	391 (27.3)	120 (43)		175(51.9)	126 (45.5)	
Surgery plus radio and chemo	400 (32.0)	65 (29)	0.001	205 (57.4)	83 (29.2)	0.460
AJCC [NO. $(\%)$]	220 (22)	EQ (17)	0.001	01 (14 0)	49 (1 (E)	0.469
1	329 (23)	50(17)		81 (14.8)	48 (16.5)	
	326 (22.8)	48(16.4)		113 (20.6) 204 (EE E)	48(16.5)	
	621 (43.4)	164 (56)		304 (55.5)	164 (56.4)	
	154 (10.8)	31 (10.6)		50 (9.1)	31 (10.6)	
Overali survival [INO. (%)]	10/2 (72.0)	225 (90.2)	0.01	202 (71 7)	222 (00.1)	0 000
1-year	1045 (72.9)	200 (ðu.2) 151 (51 5)	0.01	373 (/1./) 220 (41.8)	233 (80.1)	0.008
3-year	659 (46.1)	151 (51.5)	0.088	229 (41.8)	149 (51.2)	0.009
5-year	420 (29.4)	90 (30.7)	0.646	148 (27)	88 (30.2)	0.032

3.4. Development and Assessment of the Nomogram

To further verify that ELNs \geq 32 can improve the prognosis of patients with GSRC, a nomogram was developed to predict the survival outcome in patients with GSRC based on the optimal number of ELNs and factors selected from Cox analysis. Table 4 shows the univariate and multivariate Cox analysis of factors influencing survival outcomes. The multivariate Cox analysis results revealed that age \geq 75, size > 5 cm, ELNs < 32, tumor located in cardia, treatment with surgery only, and AJCC IV stage were associated with a poorer prognosis. The survival-prediction model of the nomogram was developed based on the above factors (Figure 3).

		Univariate Analyses	Multivariable Analyses		
Variable	n	р	HR	CI	р
Age					
<60	523 (43.4)	Reference	-	-	Reference
60–75	481 (39.9)	0.002	1.251	1.065-1.471	0.006
>75	202 (16.7)	< 0.001	1.898	1.543-2.335	< 0.001
Sex					
Male	613 (50.8)	Reference			
Female	593 (49.2)	0.268			
Race	. ,				
White	797 (66.1)	Reference			
Black	145 (12)	0.471			
Other	264 (21.9)	0.216			
Grade					
Ι	1 (0.1)	Reference			
Π	32 (2.7)	0.819			
III	1141 (94.6)	0.82			
IV	32 (2.7)	0.823			
Size					
$\leq 1 \text{ cm}$	85 (7)	Reference	-	-	Reference
1–3 cm	359 (29.8)	< 0.001	1.813	1.097-2.996	0.02
3–5 cm	316 (26.2)	< 0.001	2.312	1.382-3.868	0.001
>5 cm	446 (37)	< 0.001	3.131	1.867-5.251	< 0.001
ELNs					
<32	1004 (83.3)	Reference	-	-	Reference
\geq 32	202 (16.7)	0.014	0.633	0.518-0.775	< 0.001
Site					
Others	451 (37.4)	Reference	-	-	Reference
Cardia	165 (13.7)	0.362	1.282	1.017-1.617	0.036
Fundus/Body	177 (14.7)	0.633	1.234	0.981-1.533	0.072
Antrum/Pylorus	413 (34.2)	0.042	0.967	0.729-1.282	0.814
Therapy					
Surgery only	447 (37.1)	Reference	-	-	Reference
Surgery plus radio	14 (1.2)	0.006	1.090	0.617–1.926	0.767
Surgery plus chemo	354 (29.4)	0.012	0.619	0.508 - 0.754	< 0.001
Surgery plus radio and chemo	391 (32.4)	0.585	0.529	0.434 - 0.644	< 0.001
AJCC					
Ι	273 (22.6)	Reference	-	-	Reference
Π	268 (22.2)	< 0.001	3.127	2.308-4.235	< 0.001
III	540 (44.8)	< 0.001	6.150	4.577-8.264	< 0.001
IV	125 (10.4)	< 0.001	9.373	6.701–13.111	< 0.001

Table 4. Univariate and multivariable analyses in the derivation cohort.

Figure 4a shows that the ROC curve of the nomogram predicted the 1-, 3-, and 5-year OS rate in the training cohort with AUCs of 0.790, 0.816, and 0.782, respectively. In the testing cohort (Figure 4b), the AUCs of the 1-, 3- and 5-year OS rates were 0.763, 0.809, and 0.804, respectively, indicating that the model had good predictive capability (Table 5). Figure 5 shows the calibration curve of the nomogram for predicting the OS rate at 1, 3, and 5 years, demonstrating good agreement between prediction and observation in the primary cohort (Figure 5a–c) and testing cohort (Figure 5d–f). DCA was used to compare nomogram clinical availability with AJCC staging. As illustrated in Figure 6, the *y*-axis represents the net benefit and the *X*-axis represents the probability of survival predicted by the model.

10 Points 60-75 Age :60 -76 ELNs<32 Examined lymph nodes ELNs=32 Surgery plus Chemo Surgery plus Radio Treatment Surgery plus Radio and Chem Surgery only Ŋ AJCC ŵ 1-3cm >5cm Turnor Size 3-5cm . Icm Others Cardia Site Fundus/Body Total Points 6 à 10 14 16 18 20 22 24 26 1-year surviva 0.95 0.9 0.85 0.8 0.75 0.7 0.6 0.5 3-year survival 0.95 0.9 0.85 0.8 0.75 0.7 0.6 0.6 5-year survival 0.95 0.9 0.65 0.8 0.75 0.7 0.6 0.5

DCA graphically showed that the nomogram was superior to AJCC staging under clinical conditions and that using a nomogram provided a greater net benefit to patients.

Figure 3. Nomogram based on examined lymph nodes for predicting survival outcome in patients with gastric signet ring cell carcinoma.



Figure 4. The area under the receiver operating characteristic curve (AUROC) of the model in predicting 1-, 3-, and 5-year overall survival (OS) in the training cohort (**a**) and 1-, 3-, and 5-year OS in the validation cohort (**b**).

Table 5. Pro	ognostic effic	iency of the	nomogram of	OS.
--------------	----------------	--------------	-------------	-----

			AUC	
	C-Index (Internal Validation)	1-Year Survival	3-Year Survival	5-Year Survival
Training cohort	0.748	0.790	0.816	0.782
Testing cohort	-	0.763	0.809	0.804
XGBoost				
Including ELNs		0.803	0.832	0.806
Excluding ELNs		0.753	0.803	0.792

Table 5 and Figure 7 show predictive ability with or without the optimal number of ELNs using the XGBoost algorithm. The results showed that the XGBoost model including the optimal number of ELNs (n = 32) had better predictive ability than the XGBoost model excluding the optimal number of ELNs in predicting the OS rate at 1 (AUC: 0.803 vs. 0.753), 3 (AUC: 0.832 vs. 0.803), and 5 years (AUC: 0.806 vs. 0.792).



Figure 5. Calibration curves of the model for predicting 1- (**a**), 3- (**b**), and 5-year OS (**c**) in the training cohort and 1- (**d**), 3- (**e**), and 5-year (**f**) OS in the validation cohort.



Figure 6. Decision curve analysis for the nomogram.



Figure 7. The area under the receiver operating characteristic curve (AUROC) of the XGBoost model for predicting 1- (**a**), 3- (**b**), and 5- (**c**) year overall survival with or without the optimal number (n = 32) of examined lymph nodes (ELNs).

4. Discussion

Although the overall incidence of gastric cancer has decreased in recent decades, with the massive eradication of *Helicobacter pylori*, the incidence of GSRC is still increasing, especially in the young population [16,17]. As one of the most malignant types of GC, the prognosis for GSRC is closely associated with the stage. The prognosis for early GSRC is better than that for non-GSRC, while the prognosis for advanced GSRC is worse than for non-GSRC [18]. Furthermore, different stages have different treatment guidelines, which also induce differences in the survival prognosis of patients [2,11]. Therefore, determining accurate staging is crucial to improve the prognosis of patients with GSRC. The number of positive lymph nodes is a crucial factor for cancer staging. A high number of positive lymph nodes is also the most significant adverse prognostic indicator for GSRC because it provides a surrogate measure for both nodal burden and the likelihood of missed locoregional disease [13,19]. The number of ELNs provides information about the thoroughness of intraoperative lymph node clearance and is an accurate method for confirming the number of positive lymph nodes [20]. Adequate ELNs help determine prognosis and accurate staging, providing guidance for subsequent treatment and surveillance programs following treatment [11]. With inadequate ELNs, N1 and N2 patients may be misdiagnosed as N0 and receive different treatment and simpler monitoring procedures after treatment, resulting in a worse prognosis [1]. Furthermore, patients could have a higher risk of recurrence with inadequate ELNs, which might not cover the metastasized lymph nodes [20]. Therefore, in our study, we primarily focused on the number of ELNs and aimed to determine the optimal number of dissected lymph nodes.

According to the NCCN Clinical Practice Guidelines in Oncology (Version 2.2022), a minimum of 16 ELNs is recommended during GC surgery [11]. While it is true that patients with fewer than 16 examined lymph nodes may present a poorer prognosis, this value does not reflect the potential therapeutic value of expanded lymph node dissection for most patients [13]. Therefore, numerous studies have been carried out to identify the ELN threshold for survival advantage and oncologic benefit. For example, Smith et al. indicated that every increase of 10 lymph node dissections increases patient survival by approximately 7%, and this survival advantage can be sustained up to a threshold of 40 lymph node dissections [21]. Brenkman et al. recommended a minimum of 30 ELNs was associated with improved survival for patients with advanced-stage GC [23]. However, their recommended thresholds varied significantly, which was likely to affect clinical practice, and it is not appropriate to apply the same standards to GSRC, which is more malignant and prone to metastasis.

Currently, there are limited studies identifying the optimal number of ELNs for GSRC. Therefore, it is clinically important to investigate the optimal number of ELNs for GSRC. In our study, we analyzed patients from 2010 to 2015 in the SEER database and confirmed that 32 was the adequate cutoff number for lymph node dissection. We also confirmed that prognosis would be improved with ELNs > 32. Data from 2010 to 2015 is more comprehensive, including the availability of the complete 7th-edition AJCC classification, in addition to more complete 5-year survival data. The X-title software was used to confirm 32 as the optimal cutoff value. The chi-square values and relative risk were listed when the number of lymph nodes examined was set at 15–35. When 16 or 25 was used as the cutoff point in our modeling data, the results showed no difference in prognosis. However, when 30 was used as the cutoff value, although the *p*-value showed significance, the chi-square value and the relative risk were both less than when 32 was used as the critical value. Moreover, we compared 1-, 3-, and 5-year survival rates in groups of 32, and developed a clinical predictive model based on 32 ELNs; all of these showed that a minimum of 32 ELNs for GSRC patients could improve the survival rate. Additionally, the results of the XGBoost algorithm showed that the model including the optimal number of ELNs had better predictive ability than the model excluding the optimal number of ELNs in predicting the OS rate for patients with GSRC. Although there was little difference in the

predictive ability for 5-year survival, a significant difference was seen in the predictive ability of the XGBoost model containing the optimal number of ELNs at 1 and 3 years. This may be an issue with the data sample, and future expansion of the data sample may reveal significant differences in 5-year survival rates. Therefore, we considered the minimum of 32 ELNs to have a wider practical application for GSRC patients. Some believe that additional resection of ELNs can increase the difficulty of surgery and the potential risk of complications, resulting in a negative prognosis. However, previous studies including large samples have shown that ELNs in patients with gastric cancer \geq 30 can significantly improve survival prognosis [13,21,23–25]. Furthermore, the 8th edition of the AJCC Cancer Staging Manual highlights that the removal of \geq 30 nodes is desirable for patients with gastric cancer, which is similar to our results [26]. Therefore, we believe that additional ELNs do not worsen the prognosis of patients and would contribute to more completed tumor clearance and better staging.

Interestingly, our results are also similar to previous studies exploring the optimal number of ELNs for GC [12,13,23,27]. Among these, a large cohort study conducted by Huang et al. by analyzing two databases of patients with non-metastatic gastric adenocarcinoma found the minimum number of ELNs and the optimal number of ELNs for gastric adenocarcinoma to be 17 and 33, respectively [27]. Accordingly, we could venture a guess that the optimal number of ELNs for GC might be concentrated around that level. Therefore, we suggest that the number of ELNs should be at least greater than 32 in patients with surgically resectable GC. However, the current studies in this area are still primarily retrospective data analyses, and there is definitely a need to design higher-quality multicenter randomized controlled studies to investigate the optimal number of ELNs in GC more accurately in the future.

There are certain limitations in the current study. First, this is a retrospective study, which may include biases and potential confounding variables. Second, this study only focused on the number of ELNs and ignored the location of lymph nodes, which may be another important factor affecting the prognosis, for which a further study will be conducted in the future. Third, given the characteristics of the data relied on in our own study, the results of this study may be more applicable to patients who did not receive neoadjuvant chemotherapy due to the inability to differentiate between chemotherapy information, but this study still has reference value for other patients with GSRC treated with chemotherapy (either neoadjuvant or adjuvant). In future studies based on our own unit center, we will further differentiate chemotherapy and explore the effects of neoadjuvant, adjuvant, and neoadjuvant + adjuvant subgroups on the stability of the model to fully validate and improve the model and make the prediction results more stable. Fourth, the data for this study were obtained from authoritative public databases whose data standards were all uniform, but the data did in turn come from various subcenters. Therefore, there may be diversity in the way lymph nodes were examined and processed. However, the data entered into the database are worthy of recognition, and the results of our study still have reference value. Fifth, the selected SEER cases do not cover all states in the United States, and the applicability to Asian populations needs further study. Therefore, future validation of this study in Asian patients with GSRC is needed. However, this study included a large sample size of GSRC and was the first to investigate the optimal number of ELNs to be cleared for GSRC. Hence, this study is still of significant clinical relevance.

In conclusion, 32 was demonstrated to be the optimal number of ELNs for adequate nodal staging, and the prognostic model based on the optimal number of ELNs also showed good performance in predictive ability.

Author Contributions: Conceptualization, Y.D. and Z.L.; data curation, Y.L. and X.Y.; investigation, X.Y.; methodology, Y.D.; resources, Z.L.; supervision, W.L., J.T. and Y.D.; visualization, Z.L.; writing—original draft, Y.L.; writing—review & editing, Y.L. and J.X. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data available on request due to restrictions e.g., privacy or ethical.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Smyth, E.C.; Nilsson, M.; Grabsch, H.I.; van Grieken, N.C.; Lordick, F. Gastric cancer. *Lancet* 2020, 396, 635–648. [CrossRef] [PubMed]
- Li, Y.; Zhu, Z.; Ma, F.; Xue, L.; Tian, Y. Gastric Signet Ring Cell Carcinoma: Current Management and Future Challenges. *Cancer Manag. Res.* 2020, 12, 7973–7981. [CrossRef] [PubMed]
- 3. Arai, T. Where does signet-ring cell carcinoma come from and where does it go? *Gastric Cancer* 2019, 22, 651–652. [CrossRef] [PubMed]
- Tang, C.-T.; Chen, Y.; Zeng, C. Prognostic analysis of gastric signet ring cell carcinoma and mucinous carcinoma: A propensity score-matched study and competing risk analysis. *Aging* 2020, 12, 22059–22077. [CrossRef]
- Zhang, S.; Liu, Y.; Jiao, Z.; Li, Z.; Wang, J.; Li, C.; Qu, X.; Xu, L. Development and Validation of a Prognostic Nomogram for Gastric Signet Ring Cell Carcinoma: A Multicenter Population-Based Study. *Front. Oncol.* 2021, 11, 603031. [CrossRef]
- Yang, X.-F.; Yang, L.Y.X.-F.; Mao, X.-Y.; Wu, D.-Y.; Zhang, S.-M.; Xin, Y. Pathobiological behavior and molecular mechanism of signet ring cell carcinoma and mucinous adenocarcinoma of the stomach: A comparative study. *World J. Gastroenterol.* 2004, 10, 750–754. [CrossRef]
- Wang, H.; Peng, Y.; Huang, Q.; Wu, J.; Zhang, M. Prognostic Nomograms for Nonelderly Adults with Gastric Signet Ring Cell Carcinoma. *BioMed Res. Int.* 2021, 1274527. [CrossRef]
- Morkavuk, Ş.B.; Çulcu, S.; Tez, M.; Ünal, A.E. The efficiency of D1(+) lymphadenectomy in signet ring cell carcinoma: Comparison of postoperative early and late outcomes between standard lymphadenectomy and D1(+) lymphadenectomy. *Libyan J. Med.* 2021, 16, 1973761. [CrossRef]
- 9. Wei, F.; Lyu, H.; Wang, S.; Chu, Y.; Chen, F. Positive lymph node ratio as a novel indicator of prognosis in gastric signet ring cell carcinoma: A population-based retrospective study. *Transl. Cancer Res.* **2020**, *9*, 3658–3668. [CrossRef]
- 10. Xu, Z.; Jing, J.; Ma, G. Development and validation of prognostic nomogram based on log odds of positive lymph nodes for patients with gastric signet ring cell carcinoma. *Chin. J. Cancer Res.* **2020**, *32*, 778–793. [CrossRef]
- Ajani, J.A.; D'Amico, T.A.; Bentrem, D.J.; Chao, J.; Cooke, D.; Corvera, C.; Das, P.; Enzinger, P.C.; Enzler, T.; Fanta, P.; et al. Gastric Cancer, Version 2.2022, NCCN Clinical Practice Guidelines in Oncology. J. Natl. Compr. Cancer Netw. 2022, 20, 167–192. [CrossRef] [PubMed]
- 12. Hu, Q.; Pan, S.; Guo, Z. A novel pN3 gastric cancer staging system with superior prognostic utility based upon the examination of over 31 lymph nodes: A propensity score-matching analysis. *BMC Gastroenterol.* **2021**, *21*, 352. [CrossRef] [PubMed]
- Erstad, D.J.; Blum, M.; Estrella, J.S.; Das, P.; Minsky, B.D.; Ajani, J.A.; Mansfield, P.F.; Ikoma, N.; Badgwell, B.D. Navigating Nodal Metrics for Node-Positive Gastric Cancer in the United States: An NCDB-Based Study and Validation of AJCC Guidelines. *J. Nat. Compr. Cancer Netw.* 2021, 19, 86–97. [CrossRef] [PubMed]
- 14. Camp, R.L.; Dolled-Filhart, M.; Rimm, D.L. X-tile: A new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin. Cancer Res.* 2004, *10*, 7252–7259. [CrossRef]
- 15. Dabiri, Y.; Van der Velden, A.; Sack, K.L.; Choy, J.S.; Kassab, G.S.; Guccione, J.M. Prediction of Left Ventricular Mechanics Using Machine Learning. *Front. Phys.* **2019**, *7*, 117. [CrossRef]
- 16. Taghavi, S.; Jayarajan, S.N.; Davey, A.; Willis, A.I. Prognostic significance of signet ring gastric cancer. *J. Clin. Oncol.* **2012**, 30, 3493–3498. [CrossRef]
- 17. Li, Y.; Ma, F.-H.; Xue, L.-Y.; Tian, Y.-T. Neoadjuvant chemotherapy vs upfront surgery for gastric signet ring cell carcinoma: A retrospective, propensity score-matched study. *World J. Gastroenterol.* **2020**, *26*, 818–827. [CrossRef]
- Kerckhoffs, K.G.P.; Liu, D.H.W.; Saragoni, L.; van der Post, R.S.; Langer, R.; Bencivenga, M.; Iglesias, M.; Gallo, G.; Hewitt, L.C.; Fazzi, G.E.; et al. Mucin expression in gastric- and gastro-oesophageal signet-ring cell cancer: Results from a comprehensive literature review and a large cohort study of Caucasian and Asian gastric cancer. *Gastric Cancer* 2020, 23, 765–779. [CrossRef]
- 19. Lee, H.K.; Yang, H.-K.; Kim, W.H.; Lee, K.U.; Choe, K.J.; Kim, J.-P. Influence of the number of lymph nodes examined on staging of gastric cancer. *Br. J. Surg.* 2001, *88*, 1408–1412. [CrossRef]
- 20. Zhu, Z.; written on behalf of AME Lung Cancer Collaborative Group; Song, Z.; Jiao, W.; Mei, W.; Xu, C.; Huang, Q.; An, C.; Shi, J.; Wang, W.; et al. A large real-world cohort study of examined lymph node standards for adequate nodal staging in early non-small cell lung cancer. *Transl. Lung Cancer Res.* **2021**, *10*, 815–825. [CrossRef]
- 21. Smith, D.D.; Schwarz, R.R.; Schwarz, R.E. Impact of total lymph node count on staging and survival after gastrectomy for gastric cancer: Data from a large US-population database. *J. Clin. Oncol.* 2005, 23, 7114–7124. [CrossRef] [PubMed]

- Brenkman, H.J.F.; Goense, L.; Brosens, L.A.; Mohammad, N.H.; Vleggaar, F.P.; Ruurda, J.P.; van Hillegersberg, R. A High Lymph Node Yield is Associated with Prolonged Survival in Elderly Patients Undergoing Curative Gastrectomy for Cancer: A Dutch Population-Based Cohort Study. Ann. Surg. Oncol. 2017, 24, 2213–2223. [CrossRef]
- Ichikura, T.; Ogawa, T.; Chochi, K.; Kawabata, T.; Sugasawa, H.; Mochizuki, H. 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. [CrossRef]
- Schwarz, R.E.; Smith, D.D. Clinical impact of lymphadenectomy extent in resectable gastric cancer of advanced stage. *Ann. Surg.* Oncol. 2007, 14, 317–328. [CrossRef] [PubMed]
- Mirkin, K.A.; Hollenbeak, C.S.; Wong, J. Greater Lymph Node Retrieval Improves Survival in Node-Negative Resected Gastric Cancer in the United States. J. Gastric Cancer 2017, 17, 306–318. [CrossRef] [PubMed]
- Amin, M.B.; Greene, F.L.; Edge, S.B.; Compton, C.C.; Gershenwald, J.E.; Brookland, R.K.; Meyer, L.; Gress, D.M.; Byrd, D.R.; Winchester, D.P. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. *CA Cancer J. Clin.* 2017, 67, 93–99. [CrossRef] [PubMed]
- 27. Huang, L.; Zhang, X.; Wei, Z.; Xu, A. Importance of Examined Lymph Node Number in Accurate Staging and Enhanced Survival in Resected Gastric Adenocarcinoma—The More, the Better? A Cohort Study of 8,696 Cases From the US and China, 2010–2016. *Front. Oncol.* **2020**, *10*, 539030. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.