Received: 12 October 2018 | Revised: 18 January 2019 | Accepted: 30 January 2019

DOI: 10.1002/jso.25411

RESEARCH ARTICLE

WILEY Journal of SURGICAL ONCOLOGY

Ratio-based staging systems are better than the 7th and 8th editions of the TNM in stratifying the prognosis of gastric cancer patients: A multicenter retrospective study

Annamaria Agnes MD[1,2] | Alberto Biondi MD[1,2] | Ferdinando M Cananzi MD[3] | Stefano Rausei MD[4] | Rossella Reddavid MD[5] | Vito Laterza MD[2] | Federica Galli MD[4] | Vittorio Quagliuolo MD[3] | Maurizio Degiuli MD[5] | Domenico D'Ugo MD[1,2] | Roberto Persiani MD[1,2]

[1]Dipartimento Scienze Gastroenterologiche, Endocrino-Metaboliche e Nefro-Urologiche, Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Rome, Italy

[2]Department of Surgery, Università Cattolica del Sacro Cuore, Rome, Italy

[3]Department of Surgery, Surgical Oncology Unit, Humanitas Clinical and Research Center, Milan, Italy

[4]Department of Surgery, ASST Settelaghi, Varese, Italy

[5]Department of Oncology, Surgical Oncology, and Digestive Surgery, San Luigi University Hospital (S.L.U.H.), University of Turin, Turin, Italy

Correspondence
Alberto Biondi, MD, Dipartimento Scienze Gastroenterologiche, Endocrino-Metaboliche e Nefro-Urologiche, Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Largo Francesco Vito n. 1, 00168 Rome, Italy.
Email: alberto.biondi@policlinicogemelli.it

Abstract

Background: The current and the previous editions of the tumor-node-metastasis (TNM) system for gastric cancer (GC; TNM8 and TNM7) have a high risk of stage-migration bias when the node count after gastrectomy is suboptimal. Hence, they are possibly not the optimal staging systems for GC patients. This study aims to compare the TNM with two systems less affected by the stage-migration bias, namely, the lymph nodes ratio (LNR) and the log odds of positive lymph nodes (LODDS), to assess which one is the best in stratifying the prognosis of GC patients.

Methods: The sample study included 1221 GC patients. Two 7-cluster staging systems based on the combination of pT categories and LNR and LODDS categories (TLNR and TLODDS) were compared with the two last editions of TNM, using the Akaike information criteria, the Bayesian information criteria, and the receiver operating characteristic (ROC) curve graphs. Further validation on an independent sample of 251 patients was carried out.

Results: The univariable and multivariable analyses and the ROC curves detected an advantage of the TLNR and TLODDS systems over the TNM. The TLNR and TLODDS showed the best accuracy both in the subgroup of patients with ≥16 nodes examined. The results were confirmed in the validation analysis.

Conclusions: TLNR and TLODDS staging systems should be considered a valid implementation of the TNM for the prognostic stratification of GC patients. If these results are confirmed in further studies, the future implementation of the TNM should consider the introduction of the LNR or the LODDS along with the number of metastatic nodes.

KEYWORDS
gastric cancer, lymph nodes ratio (LNR), staging system, tumor-node-metastasis (TNM)

1 | INTRODUCTION

Gastric cancer (GC) is the fifth most frequent malignancy and the third leading cause of cancer death worldwide.[1] The prognosis of GC is influenced by numerous factors. Among these, the nodal involvement is considered one of the most influent.[2,3] For this reason, the prognostic accuracy of the node subsection (N) of the American Joint Committee on Cancer (AJCC) tumor-node-metastasis (TNM) classification has been pursued in all its late editions. Although the TNM 8th edition proved to be superior to the TNM 7th edition in predicting overall survival,[4-7] both editions are subject to the risk of stage migration. The phenomenon of stage migration occurs when the number of examined lymph nodes is insufficient. In particular, when it is less than 16, N3b patients may be improperly classified as N3a, and understaging may occur.[8] One of the solutions proposed to overcome the potential bias associated with an inadequate number of node examined and with stage migration is the use of the lymph nodes ratio (LNR), namely, the ratio between the number of metastatic nodes and the total examined nodes, as an alternative to the N classification. The LNR accounts for the total number of examined nodes, and offers information on the extent of the node dissection, on the accuracy of the pathologic staging and, possibly, on the immune status of the patients.[2,9] The LNR carries more information than the N classification, which is based on the absolute number of metastatic nodes, and therefore it has been investigated as a possible alternative to the N classification both in patients with an inadequate and with an adequate node count in previous studies. Many of these studies have shown a better predictive value of LNR when compared with the N classification.[10,11] Others, however, did not confirm this advantage.[12] The LNR, which weights for the total number of nodes examined, has the same limit of the N classification in not being able to stratify the prognosis of patients with all metastatic (LNR = 1) or no metastatic (LNR = 0) nodes. For this reason, another solution proposed to minimize the limits of the N and LNR classifications was the use of the log odds of positive lymph nodes (LODDS), namely, the logarithm of the ratio between metastatic nodes and non metastatic nodes,[13] which carries all the adjunctive information and the advantages of the LNR system and further discriminates among patients with edge LNR values (LNR 0 to 1).[13,14]

To date, the prognostic advantage of alternative staging systems has been investigated in a significant number of studies, which mostly compared the difference between the sole nodal classification systems (N, LNR e LODDS), with discordant results.[12,14-22] Most of these studies have been conducted in Asia[15-18,20-23] and, so far, no study has compared the prognostic value of a full TNM classification system based on the LNR and LODDS classification with the TNM 8th edition. Therefore, the aim of this multicenter study is to define two full classification systems based on the LNR and LODDS classification and to compare their prognostic value with those of TNM 7th and 8th editions.

2 | MATERIALS AND METHODS

2.1 | Patients

Between September 1987 and December 2015, 1221 patients with GC underwent gastrectomy in the General Surgery Unit of the Fondazione Policlinico Universitario "A. Gemelli" (Rome, Italy), in the Surgical Oncology Unit of the Humanitas Clinical and Research Center (Milan, Italy), and in the Ospedale di Circolo "Fondazione Macchi" (Varese, Italy). Patients records had been collected in three institutional databases, which were retrospectively reviewed for the purpose of this study. We selected for inclusion in this study all stage I to III patients with Siewert III and gastric carcinoma who underwent a curative-intent gastrectomy. Patients who underwent neoadjuvant therapy, patients with remnant GC, patients undergoing a R2 resection, patients in stage IV (metastatic at diagnosis), patients who died in the postoperative period, patients with incomplete data on the pathological staging and patients lost to follow-up were excluded from the study. Patients were classified by the 7th and the 8th edition of the AJCC.[24]

To integrate the results of the study, we planned an external validation analysis with an independent sample. Patients were selected from an initial population of 251 patients who underwent gastrectomy in Surgical Oncology and Digestive Surgery Unit of the San Luigi University Hospital, Orbassano, Turin, Italy, between April 1998 and December 2016. Data had been prospectively collected in an Institutional database. A total of 183 patients were selected according to the same inclusion and exclusion criteria of the principal study. All the selected patients underwent a curative-aim R0 gastrectomy with at least a D2 gastrectomy.

2.2 | Data Collection

Information collected from the medical records included: age, sex, tumor location, invasion of the gastroesophageal junction, Lauren histological type, type of gastrectomy, additional organ resection, type of lymphadenectomy, resection status, pTNM, number and location of the metastatic nodes, total examined nodes, administration of adjuvant therapy, and cancer-specific outcomes.

2.3 | Surgery + /− adjuvant therapy

Curative-aim gastrectomy consisted of a subtotal or total gastrectomy, associated with a D1, a D2 or a D2 plus lymphadenectomy. Total gastrectomy was performed for tumors involving the proximal body, fundus, and cardia of the stomach, and for distal tumors with a positive resection margin at frozen section. Multivisceral resections included splenectomy, pancreatectomy, colic or transverse mesocolic resection, hepatic resection, cholecystectomy, and appendectomy. Postoperative chemotherapy and radiochemotherapy were administered according to the local hospitals' protocols. Follow-up was planned at least every 6 months for the 2 years after gastrectomy and yearly thereafter.

The aim of this study was to identify the staging system best associated with disease-specific survival (DSS) in a population of GC patients treated by means of curative-aim gastrectomy.

2.4 | Statistics and criteria for assessing the best staging system

LNR was defined as the ratio between the number of metastatic nodes and the total number of examined nodes. Five LNR categories were defined on the basis of the current literature,[9,11,21] as LNR0 when LNR = 0, LNR1 when 0.01-0.1, LNR2 when 0.11-0.25, LNR3 when 0.26-0.40, LNR4 when >0.40. LODDS was calculated as log $\frac{(pnod+0.5)}{(tnod-pnod+0.5)}$, with *pnod* being the number of metastatic nodes and *tnod* being the number of total nodes examined. LODDS categories were defined, following the classification proposed by Sun et al:[13] LODDS1 when LODDS < −1.5; LODDS2 when − 1.5 ≤ LODDS < −1.0; LODDS3 when − 1.0 ≤ LODDS < 0.5; LODDS4 when 0.5 ≤ LODDS < 0; and LODDS5 when LODDS ≥ 0.[13]

DSS was calculated from the time of surgery. The association of every variable with DSS was tested in a univariable Cox proportional hazards model, and variables significant at 0.05 were included in a subsequent multivariable Cox proportional hazards models. Three multivariable Cox proportional hazards models were used to identify the nodal classification best associated with DSS, using respectively the number of metastatic nodes, LNR, LODDS (as continuous variables). The −2log-likelihood (−2LLH) from every multivariable analysis was used to calculate the Akaike information criteria (AIC) and Bayesian information criteria (BIC). AIC was calculated as −2LLH + 2(*df*) with df as degrees of freedom and BIC was calculated as −2LLH + (*df*) (LN(sample size]). The best predictive models were identified as the ones with smallest AICs and BICs.

Combining the pT categories with the LNR and the LODDS categories, 25 TLNR and 25 TLODDS classes were respectively identified. These classes were aggregated in 7 TLNR and 7 TLODDS clusters based on the hazard ratio (HR) for DSS.

Survival curves for every staging system were estimated with the Kaplan-Meier method. To compare the accuracy of the different staging systems, receiver operating characteristic (ROC) curve graphs and their areas under the curve (AUC) were generated, using the 5-year DSS as the measure. AUCs of the different staging systems were compared using DeLong's method. To identify the staging system best associated with DSS, four multivariable Cox proportional hazards model regressions were used. The first model included the TMN7, the second included the TNM8, the third included the TLNR clusters, and the fourth included the TLODDS clusters. AIC and BIC were calculated for every model. The best predictive models were identified as the one with the smallest AIC and BIC.

Then, a subgroup analysis was conducted for patients with <16 nodes and ≥16 nodes examined. In every subgroup, the accuracy of the four staging systems was compared using the ROC curve graphs and their AUCs with 95% confidence interval (CI), and the predictive capacity of the four staging systems was evaluated conducting a Cox univariable analysis, and calculating AIC and BIC for every model. Another subgroup analysis was planned for N0 patients, and the accuracy of the four staging systems was compared using the ROC curve graphs and their AUCs with 95% CI.

Lastly, the site-specific stage migration phenomenon for every staging system (TNM7, TNM8, TLNR, TLODDS) was expressed as the percentage of patients in which there was a stage change when considering the nodes of all the stations included in a D2 gastrectomy instead of considering the nodes that would have been included in a D1 gastrectomy (stations 1; 3-7 for subtotal gastrectomies and 1-7 for total gastrectomies). The best staging system was considered the one with the smallest percentage of stage migration.

Statistical analysis was conducted using the SPSS v.22 for Windows XP (SPSS, Chicago, IL) and MedCalc statistical software version 18.2.1 (MedCalc Software bvba, Ostend, Belgium; http://www.medcalc.org; 2018). All statistical tests were two-sided with significance set at $P \leq 0.05$.

2.5 | External validation analysis

As a validation analysis, four multivariable Cox proportional hazards model regressions (using respectively the TNM7, the TNM8, the TLNR clusters, and the TLODDS clusters) were used to calculate the AIC and BIC for every model. The best predictive model and the best staging system was identified as the one with the smallest AIC and BIC. Then, the accuracy of the different staging systems was assessed by the use of ROC curve graphs and their AUC with the 5-year DSS as the measure.

2.6 | Ethical approval

The Institutional review board Ethics Committee approved this retrospective study. The manuscript does not include any potentially identifiable patient images or data.

3 | RESULTS

3.1 | Whole population of the study

The population of this study consists of 820 patients. Their clinicopathological characteristics are presented in Table 1. The median follow-up of survivors was 75 months. The 1-year DSS was 94.1%. The 3-year DSS was 76.0%. The 5-year DSS was 63.9%.

The multivariable analyses including alternatively the number of metastatic nodes, the LNR and the LODDS (Table S1s) demonstrated an AIC of 2349 and a BIC of 2423 for the analysis including number of metastatic nodes; an AIC of 2319 and a BIC of 2392 for the analysis including the LNR; an AIC of 2319 and a BIC of 2393 for the analysis including the LODDS (Table 2).

After the clustering of the different associations between pT and the categories of LNR and LODDS, two new classification systems (TLNR, TLODDS) were established (Figures 1 and 2; Table S2s). Kaplan-Meier curves were built for every staging system (TNM7, TNM8, TLNR, and TLODDS; Figure 3). The univariable analyses conducted for every staging system (n = 820 patients) detected an AIC of 2730 and a BIC of 2758 for the TNM7; an AIC of 2726 and a BIC of 2754 for the TNM8, an AIC of

TABLE 1 Clinicopathological features of 820 patients undergoing curative-aim gastrectomy

Variables	N (%)
Sex	
M	494 (60.2)
F	326 (39.8)
Age (mean ± SD)	67 ± 13
Location	
Upper	123 (15)
Middle	290 (35.4)
Lower	395 (48.2)
Whole stomach	11 (1.3)
NA	1 (0.1)
EGJ involvement	
Yes	23 (2.8)
No	797 (97.2)
Type of gastrectomy	
Total	251 (30.6)
Subtotal	569 (69.4)
Multivisceral resection	
Yes	99 (12.1)
No	720 (87.8)
NA	1 (0.1)
Type of lymphadenectomy	
D1	275 (33.5)
D2/+	543 (66.2)
NA	2 (0.2)
Resection status	
R0	756 (92.2)
R1	58 (7.1)
NA	6 (0.7)
Adjuvant therapy	
Yes	268 (32.7)
No	497 (60.6)
NA	55 (6.7)
Lauren histological type	
Intestinal	415 (50.6)
Diffuse	268 (32.7)
Mixed	68 (8.3)
NA	69 (8.4)
7th-8th AJCC/UICC TNM classification T staging	
T0	31 (3.8)
T1	191 (23.3)
T2	131 (16.6)
T3	245 (29.9)
T4a	203 (24.8)
T4b	19 (2.3)
7th-8th AJCC/UICC TNM classification N staging	
N0	337 (41.1)
N1	156 (19.0)
N2	138 (16.8)
N3a	118 (14.4)
N3b	71 (8.7)
7th AJCC/UICC TNM classification stage grouping	
0/IA	172 (21.0)
IB	98 (12)
IIA	104 (12.7)
IIB	121 (14.8)
IIIA	102 (12.4)
IIIB	118 (14.4)
IIIC	105 (12.8)
8th AJCC/UICC TNM classification stage grouping	
0/IA	172 (21)
IB	98 (12)
IIA	104 (12.7)
IIB	121 (14.8)
IIIA	146 (17.8)
IIIB	105 (12.8)
IIIC	74 (9)
Positive nodes (mean ± SD; range)	4.48 ± 7.30; 0-52
Total nodes (mean ± SD; range)	28.06 ± 14.86; 0-86
LNR (mean ± SD; range)	0.15 ± 0.22; 0-0.97
LODDS (mean ± SD; range)	−0.99 ± 0.70; −2.2-1.4
1-y survival (N = 779)	94.1%
3-y survival (N = 684)	76%
5-y survival (N = 592)	63.9%

Abbreviations: AJCC, American Joint Cancer Commission; LNR, lymph nodes ratio; LODDS, log odds of positive lymph nodes; NA, not available; SD, standard deviation; TNM, tumor-node-metastasis; UICC, Union International Contra Cancer (International Union Against Cancer).

2682 and a BIC of 2710 for the TLNR; and an AIC of 2690 and a BIC of 2718 for the TLODDS. The multivariable analyses (n = 740 patients) detected an AIC of 2343 and a BIC of 2421 for the TNM7; an AIC of 2334 and a BIC of 2412 for the TNM8, an AIC of 2296 and a BIC of 2374 for the TLNR; an AIC of 2306 and a BIC of 2384 for the TLODDS (Table 2). The evaluation of the accuracy of the different staging system through the ROC curves (n = 592) demonstrated an AUC (95% CI) of 0.831 (0.796-0.866) for the TNM7, of 0.828 (0.793-0.863) for the TNM8, of 0.853 (0.820-0.886) for the TLNR and of 0.850 (0.817-0.883) for the TLODDS (Figure 4). The comparison between ROC curves documented a significant difference between AUCs for the TNM7 and the TLNR and TLODDS ($P = 0.0006$ and $P = 0.0090$, respectively), and between AUCs for the TNM8 and the TLNR and TLODDS ($P = 0.0002$ and $P = 0.0031$, respectively). No significant difference was detected for the comparison

TABLE 2 Akaike information criteria (AIC) and Bayesian information criteria (BIC) derived from the univariable and multivariable analysis for disease-specific survival

	−2LLH	DF	AIC	BIC	Sample size
Multivariable analysis for all patients					
N	2317	16	2349	2423	740
LNR	2287	16	2319	2392	740
LODDS	2287	16	2319	2393	740
Univariable analysis for all patients					
TNM7	2718	6	2730	2758	820
TNM8	2714	6	2726	2754	820
TLNR	2670	6	2682	2710	820
TLODDS	2678	6	2690	2718	820
Multivariable analysis for all patients					
TNM7	2309	17	2343	2421	740
TNM8	2300	17	2334	2412	740
TLNR	2262	17	2296	2374	740
TLODDS	2272	17	2306	2384	740
Univariable analysis for patients with <16 nodes examined					
TNM7	406	6	418	437	163
TNM8	405	6	417	435	163
TLNR	399	6	411	429	163
TLODDS	402	6	414	433	163
Multivariable analysis for patients with <16 nodes examined					
TNM7	384	11	406	440	160
TNM8	435	11	457	491	160
TLNR	377	11	399	433	160
TLODDS	383	11	405	439	160
Univariable analysis for patients with ≥16 nodes examined					
TNM7	2048	6	2060	2087	654
TNM8	2038	6	2050	2077	654
TLNR	2022	6	2034	2061	654
TLODDS	2028	6	2040	2067	654
Multivariable analysis for patients with ≥16 nodes examined					
TNM7	1735	16	1767	1837	597
TNM8	1722	16	1754	1825	597
TLNR	1711	16	1743	1814	597
TLODDS	1716	16	1748	1818	597
Univariable analysis for external validation					
TNM7	289	6	301	320	183
TNM8	289	6	301	321	183
TLNR	281	6	293	312	183
TLODDS	284	6	296	315	183
Multivariable analysis for external validation					
TNM7	277	7	291	313	183
TNM8	276	7	290	312	183
TLNR	273	7	287	309	183
TLODDS	271	7	285	307	183

Abbreviations: −2LLH, −2log-likelihood; DF degrees of freedom; N, number of examined lymph nodes; LNR, lymph nodes ratio; LODDS log odds of positive lymph nodes; TNM, tumor-node-metastasis.

between AUCs for the TNM7 and the TNM8 ($P = 0.4427$) and for the comparison between TLNR and TLODDS ($P = 0.3742$). Data on the location of the metastatic nodes were available for 515 patients. In these patients, the rate of site-specific stage migration was 5% for the TNM7, 5% for the TNM8, 4.7% for the TLNR, and 8.2% for the TLODDS.

3.2 | Patients with <16 nodes examined

The univariable Cox regression (n = 163 patients) showed an AIC and a BIC of 418 and 437 for the TNM7, 417 and 435 for the TNM8, 411 and 429 for the TLNR, and 414 and 433 for the TLODDS. The multivariable analyses (n = 160 patients) detected an AIC of 406 and a BIC of 440 for the TNM7; an AIC of 457 and a BIC of 491 for the TNM8, an AIC of 399 and a BIC of 433 for the TLNR; an AIC of 405 and a BIC of 439 for the TLODDS (Table 2). The evaluation of the accuracy of the different staging system through the ROC curves (n = 120 patients) demonstrated an AUC (95% CI) of 0.798 (0.709-0.887) for the TNM7, 0.793 (0.704-0.883) for the TNM8, 0.827 (0.744-0.911) for the TLNR, and 0.818 (0.734-0.902) for the TLODDS (Figure 4). The comparison between ROC curves documented a significant difference between AUCs for the TNM7 and the TLNR ($P = 0.0049$), and between AUCs for the TNM8 and the TLNR ($P = 0.0028$). No significant difference was detected for the comparison between AUCs for the TNM7 and the TNM8 ($P = 0.1158$), the TNM7 and the TLODDS ($P = 0.0939$), the TNM8 and the TLODDS ($P = 0.0522$, nearly significant), and the TLNR and TLODDS ($P = 0,2517$).

3.3 | Patients with ≥16 nodes examined

The univariable Cox regression (n = 634 patients) showed an AIC and a BIC of 2060 and 2087 for the TNM7, 2050 and 2077 for the TNM8, 2034 and 2061 for the TLNR, and 2040 and 2067 for the TLODDS. The multivariable analyses (n = 597 patients) detected an AIC of 1767 and a BIC of 1837 for the TNM7; an AIC of 1754 and a BIC of 1825 for the TNM8, an AIC of 1743 and a BIC of 1814 for the TLNR; an AIC of 1748 and a BIC of 1818 for the TLODDS (Table 2). The evaluation of the accuracy of the different staging system through the ROC curves (n = 472 patients) demonstrated an AUC (95% CI) of 0.852 (0.816-0.888) for the TNM7, 0.850 (0.814-0.886) for the TNM8, 0.862 (0.828-0.897) for the TLNR, and 0.861 (0.827-0.895) for the TLODDS (Figure 4). The comparison between ROC curves documented a nonsignificant difference between AUCs for the TNM7 and the TLNR ($P = 0.1161$), the TNM7 and the TLODDS ($P = 0.2280$), the TNM8 and the TLNR ($P = 0.0790$), the TNM8 and the TLODDS ($P = 0.1706$), the TNM7 and the TNM8 ($P = 0.8131$) and the TLNR and the TLODDS ($P = 0.6909$).

3.4 | Patients with N0 disease

The evaluation of the accuracy of the different staging system through the ROC curves (n = 230 patients) demonstrated an AUC (95% CI) of 0.678 (0.574-0.782) for the TNM7 and TNM8, of 0.691 (0.587-0.795) for the TLNR and of 0.688 (0.584-0.792) for the TLODDS (Figure 4). The comparison between ROC curves documented a nonsignificant difference between AUCs for the TNM7 and the TLNR ($P = 0.1480$), the TNM7 and the TLODDS ($P = 0.2615$), the TNM8 and the TLNR ($P = 0.1480$), the TNM8 and the TLODDS ($P = 0.2615$), the TNM7 and the TNM8 ($P = 1$), and the TLNR and the TLODDS ($P = 0,3152$).

3.5 | External validation analysis

The univariable Cox regressions (n = 183 patients) showed an AIC and a BIC of 301 and 320 for the TNM7, 301 and 321 for the TNM8, 293 and 312 for the TLNR, and 296 and 315 for the TLODDS. The multivariable analyses (n = 183 patients; variables included: age, staging system)

	N0	N1	N2	N3a, N3b
T1	IA	IB	IIA	IIB
T2	IB	IIA	IIB	IIIA
T3	IIA	IIB	IIIA	IIIB
T4a	IIB	IIIA	IIIB	IIIC
T4b	IIIB	IIIB	IIIC	IIIC

	LNR0	LNR1	LNR2	LNR3	LNR4
T1	TLNR1	TLNR1	TLNR1	TLNR1	TLNR5
T2	TLNR2	TLNR2	TLNR3	TLNR3	TLNR6
T3	TLNR2	TLNR3	TLNR4	TLNR6	TLNR6
T4a	TLNR4	TLNR5	TLNR5	TLNR6	TLNR7
T4b	TLNR6	TLNR6	TLNR6	TLNR7	TLNR7

	N0	N1	N2	N3a	N3b
T1	IA	IB	IIA	IIB	IIIB
T2	IB	IIA	IIB	IIIA	IIIB
T3	IIA	IIB	IIIA	IIIB	IIIC
T4a	IIB	IIIA	IIIA	IIIB	IIIC
T4b	IIIA	IIIB	IIIB	IIIC	IIIC

	LODDS1	LODDS2	LODDS3	LODDS4	LODDS5
T1	TLODDS1	TLODDS1	TLODDS1	TLODDS1	TLODDS4
T2	TLODDS2	TLODDS2	TLODDS2	TLODDS4	TLODDS5
T3	TLODDS2	TLODDS3	TLODDS3	TLODDS6	TLODDS6
T4a	TLODDS3	TLODDS5	TLODDS5	TLODDS6	TLODDS7
T4b	TLODDS6	TLODDS6	TLODDS6	TLODDS7	TLODDS7

FIGURE 1 Stage grouping for the TNM7, the TNM8, and the novel TLNR and TLODDS staging systems: TNM7 (upper left panel), TNM8 (lower left panel), TLNR (upper right panel), TLODDS (lower right panel). LNR, lymph nodes ratio; LODDS, log odds of positive lymph nodes [Color figure can be viewed at wileyonlinelibrary.com]

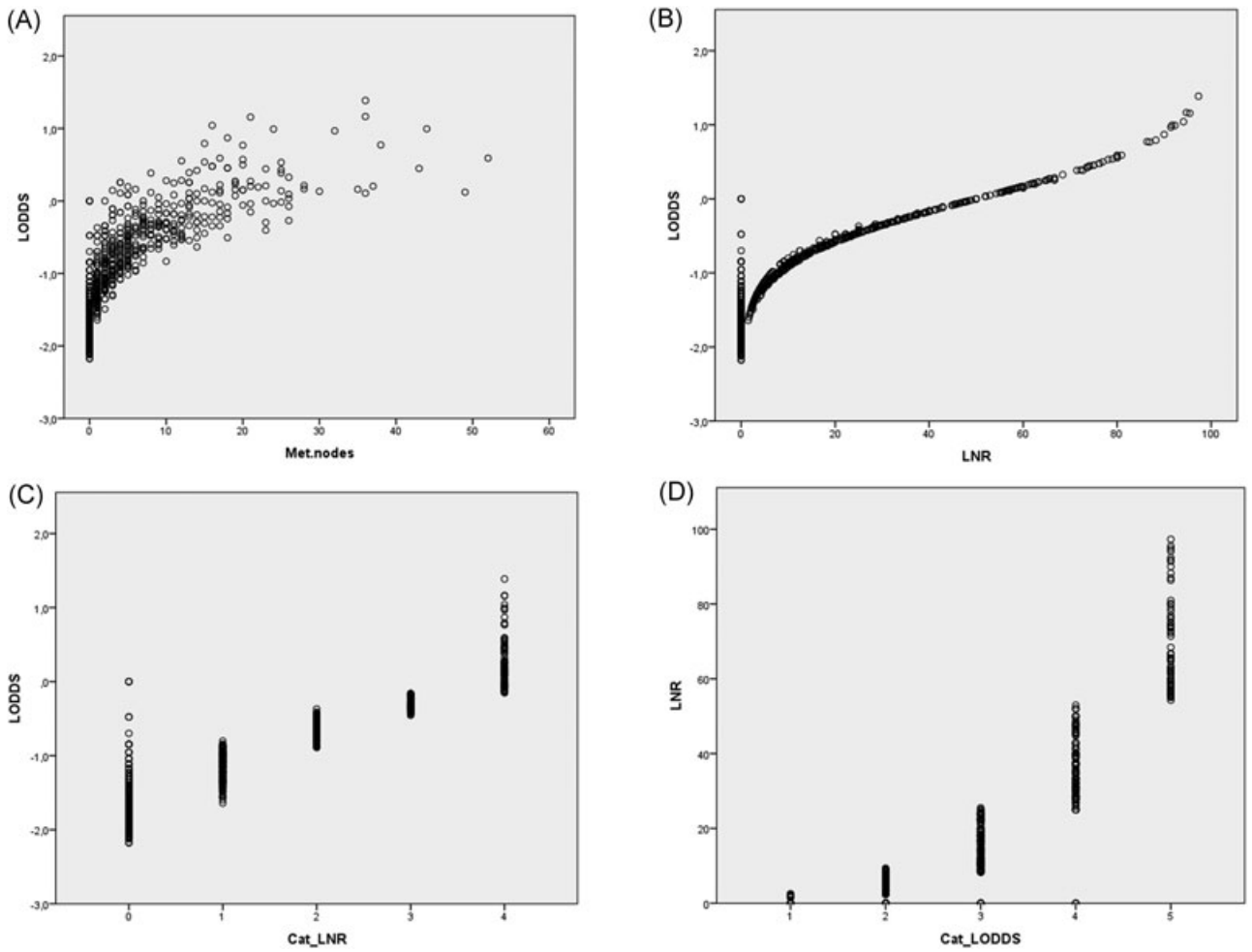

FIGURE 2 Depiction of the distribution of the LODDS and the number of lymph node metastases (A), of the LODDS and the LNR (B); of the LODDS compared with the LNR categories (C), of the LNR compared with the LODDs categories (D). LNR, lymph nodes ratio; LODDS, log odds of positive lymph nodes

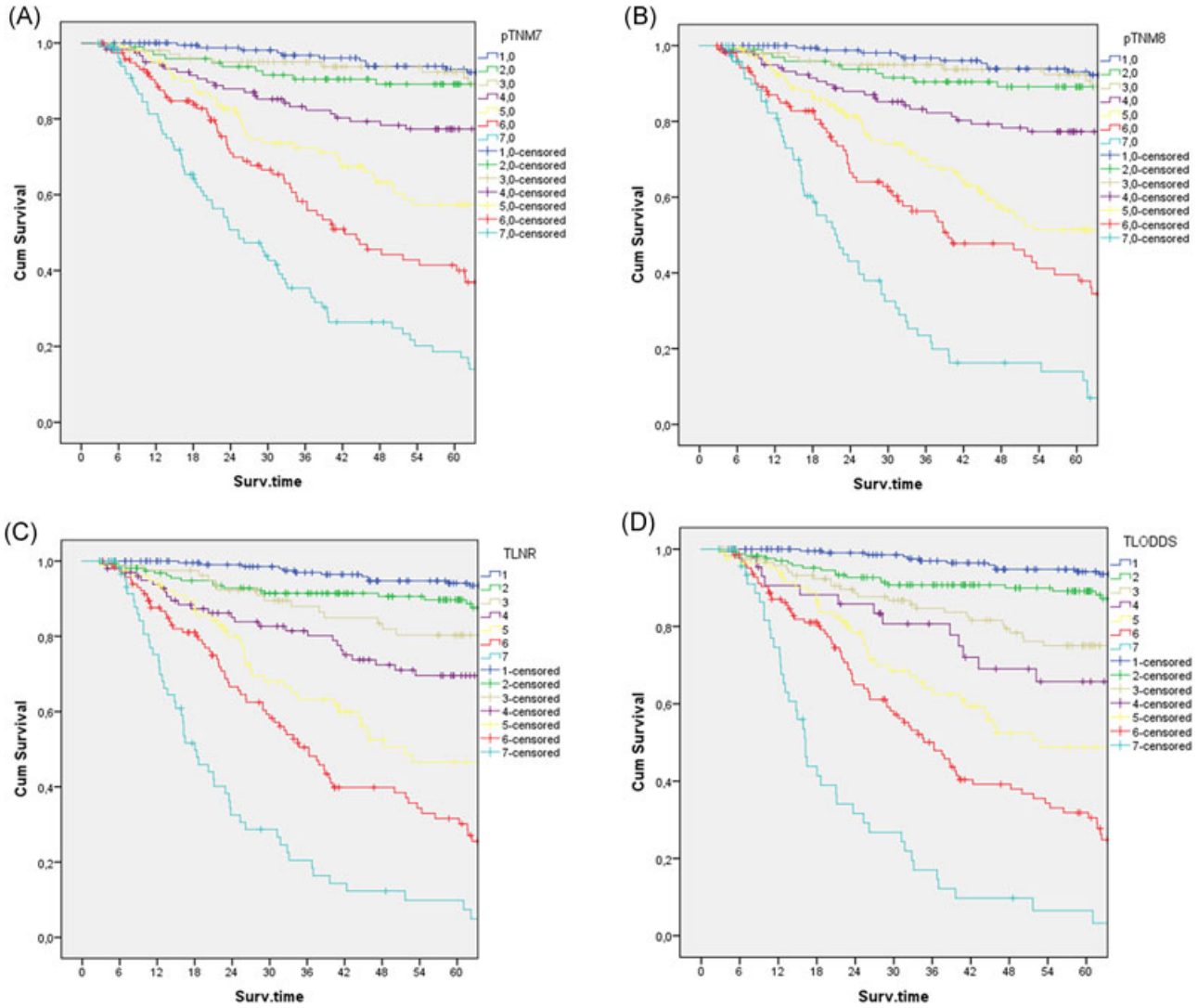

FIGURE 3 Kaplan-Meier curves showing the DSS by the TNM7 (A), the TNM8 (B), the TLNR (C), and the TLODDS (D) staging systems. DSS, disease-specific survival; LNR, lymph nodes ratio; LODDS, log odds of positive lymph nodes; TNM, tumor-node-metastasis [Color figure can be viewed at wileyonlinelibrary.com]

detected an AIC of 291 and a BIC of 313 for the TNM7; an AIC of 290 and a BIC of 312 for the TNM8, an AIC of 287 and a BIC of 309 for the TLNR; an AIC of 285 and a BIC of 307 for the TLODDS (Table 2). The evaluation of the accuracy of the different staging systems through the ROC curves (n = 107 patients) demonstrated an AUC (95% CI) of 0.818 (0.727-0.910) for the TNM7 and of 0.818 (0.726-0.909) for the TNM8, of 0.840 (0.750-0.930) for the TLNR and of 0.823 (0.728-0.917) for the TLODDS. The comparison between ROC curves documented a nonsignificant difference between AUCs for the TNM7 and the TLNR ($P = 0.2992$), the TNM7 and the TLODDS ($P = 0.8739$), the TNM8 and the TLNR ($P = 0.2736$), the TNM8 and the TLODDS ($P = 0.8482$), the TNM7 and the TNM8 ($P = 0.8912$), and the TLNR and the TLODDS ($P = 0.1955$).

4 | DISCUSSION

In this study, the LNR and LODDS proved to be superior to the number of metastatic nodes in stratifying the prognosis on GC patients. As well, the new staging systems (TLNR and TLODDS) were superior to the TNM7 and TNM8. These results were confirmed both in the subgroup of patients with <16 nodes examined and ≥16 nodes examined, and in the external validation analysis, especially in the population with <16 nodes examined, where the difference between AUCs reached the statistical significance. When LNR and LODDS were used as continuous variables, the LODDS proved to almost equivalent to the LNR, while after the categorization based on the data of the current literature, the TLNR classification showed a better predictive value and a better accuracy for the stratification of the prognosis.

The reason for the TLNR and TLODDS systems being better predictors of the DSS probably lies in the nature of these indexes. Indeed, LNR and LODDS weight for the total node count. The total node count may be an indirect indicator of the extent of the lymph node dissection.[2] Moreover, it could be an indirect indicator of the immune status of the single patient, as the total node counts is

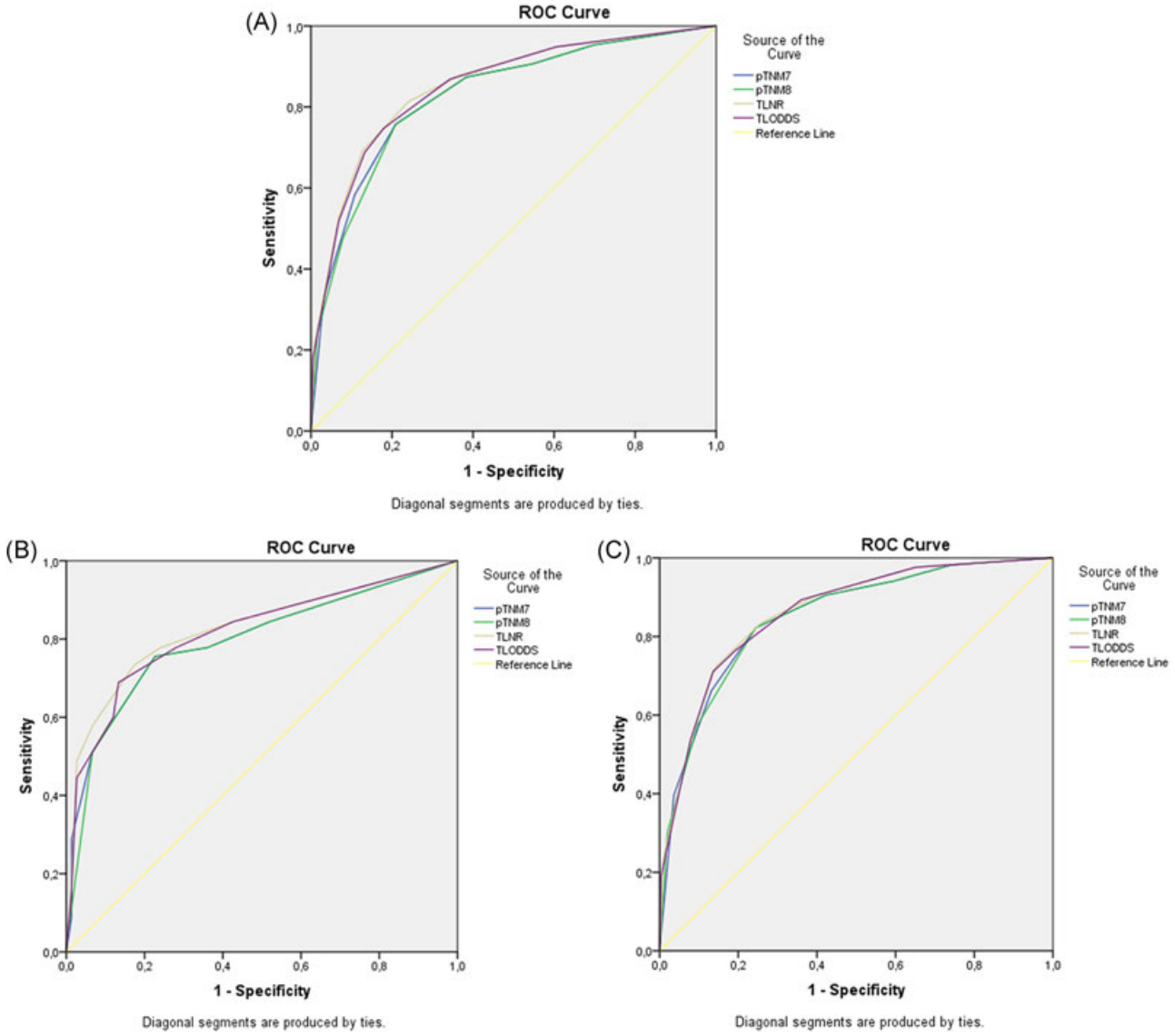

FIGURE 4 Comparison of ROC curves with TNM7, TNM8, TLNR and TLODDS staging systems (A) in all patients, (B) in patients with <16 lymph nodes examined, and (C) in patients with ≥16 nodes examined. LNR, lymph nodes ratio; LODDS, log odds of positive lymph nodes; ROC, XXX; TNM, tumor-node-metastasis [Color figure can be viewed at wileyonlinelibrary.com]

generally increased in patients with better performance status and a more efficient immune system.[9] Last, the total node count may represent an indirect indicator of the accuracy of the pathological node harvesting.[25] The importance of having an index which weights for the surgical and pathological quality is especially relevant to limit the stage migration phenomenon that may occur when the number of nodes examined is <16. The LODDS system is theoretically less dependent on the total node count (as it is related to the negative node count). This should give a LODDS based staging system a clinical advantage over an LNR-based system, as it theoretically has the advantage of identifying subgroups with different prognosis among patients with null or limited nodal involvement (ie, N0 or N1; Figure 2).[13] Hence, LODDS may identify high-risk patients (otherwise classified as N0 by TNM or LNR0 by TLNR stage) deserving closer follow-up or adjuvant therapy. However, this theoretical advantage was neither confirmed in the multivariable analysis including LNR and LODDS as continuous variables, nor after the LNR and LODDS categorization, nor in the subgroup analysis conducted on N0 patients. Overall, the TLNR system showed the best predictive value for DSS, the best accuracy for the 5-year survival and the best control over the site-specific stage migration phenomenon. The TLODDS system demonstrated an advantage over the TNM7 and TNM8 and a disadvantage compared with the TLNR. In particular, the TLODDS system showed to be the most affected by the site-specific stage migration phenomenon. This is probably due to the mathematical properties of LODDS, as the denominator is represented by the number of negative nodes and therefore the ratio between the positive nodes (that represent a relatively stable number in the D1 setting) and the negative nodes it is more influenced by the removal of D2 stations when compared to LNR, whose denominator is the total number of nodes examined (a greater number that represents a more

stable denominator). Another possible explanation for the unsatisfactory performance of the LODDS in this population is the fact that there are no patients with 100% of metastatic lymph nodes (Figure 2), and therefore the theoretical advantage of the LODDS could apply only to the N0/LNR0 category. In this category, LODDS should be able to identify high- and low-risk N0/LNR0 patients. However, as it is clear from the results, such a distinction does not seem to be so meaningful from a prognostic standpoint. Instead, the strong discrimination among patients with a different burden of node involvement in advanced stages makes LNR able to overcome the limits of the TNM7 and TNM8 and to better predict the prognosis. In Figure 2, it is shown how the LODDS system has a better discriminating capability in the earlier stages of node involvement when compared with the LNR and how the LNR system has a better discriminating capability in the later stages.

Our results are in line with those of previous authors. Lee et al[21] proved the LNR to be superior to the conventional nodal staging systems (pN from the TNM7). Sun et al[13] demonstrated that the use of LODDS categories is more reliable than the use of pN and LNR categories in stratifying the prognosis of GC patients. Spolverato et al[14] also found the LODDS to be the best predictor of survival when the lymph node status was assessed as a continuous variable, but when the lymph node status was categorized the LNR became the best predictor of survival. This last finding was confirmed by Liu et al,[17] and is consistent with the results of this study, where LODDS was an almost equivalent predictor to LNR when used as a continuous variable, while once categorized and clustered in the TLODDS staging systems it lost part of its predictive value compared with LNR. These results highlight the limits of a categorical classification and the possible advantages of a staging system that use continuous variables. However, such a system would be more difficult to design and to apply, therefore a possible solution could be the identification of new ideal cutoffs for LNR and LODDS on a big sample collected through a multicenter design.

Full staging systems based on LNR and LODDS categories have been compared with the TNM only in a few studies. Qiu et al[18] compared the TNM7 with two staging systems based on LNR and LODDS, identifying the TLODDS system as the best performing one for predicting survival. Wang et al,[22] instead, compared the TNM7 with a TLNR system, demonstrating the system based on TLNR to be more accurate than the TNM7.

This study has some limitations. The first is due to the fact that it is not international, being based only on the Italian population. Our results should probably be validated with populations from other countries to further strengthen the results. The limited number of patients in the validation cohort is another limitation of the study. However, the results were consistent with the main analysis.

This study has also many strengths, as it is the first comparing the full TLNR and TLODDS staging systems with the latest version of the TNM. Moreover, it is based on a multicenter design and on a consistent number of patients, and it has the adjunctive strength of being based on an extensive and accurate statistical validation of the new models. Our results indicate a clear advantage of the TLNR system and the TLODDS system over the TNM7 and the TNM8, especially in the population of patients with <16 nodes examined.

The implications of this study are timely. Indeed, it is of great importance to identify an optimal staging system to define the best treatment strategy and the best surveillance scheme for every stage of the disease. Currently, many novel extra-TNM factors are available to further characterize the disease of every patient with GC.[26-28] However, as these novel factors have not yet been included in the current staging systems, it is still fundamental to ensure the availability of a system which is the most accurate and reliable as possible. This system should be solid in every subgroup of patients, even in those where a poor primary lymph node count or a suboptimal surgical treatment or inaccurate pathological retrieval of lymph nodes could compromise a correct staging.

5 | CONCLUSION

In this study, the staging systems based on the rate between node counts proved to be more reliable and accurate than the TNM7 and TNM8 in the prognostic stratification of GC patients. The prognostic stratification advantage was consistent in the subgroup of patients with <16 examined nodes and in the subgroup of patients with ≥16 nodes examined, and it was confirmed by the external validation analysis. TLNR and TLODDS staging systems should be considered a valid implementation of the TNM for the prognostic stratification of GC patients. If these results are confirmed in further studies, the future implementation of the TNM should consider the introduction of the LNR or the LODDS along with the number of metastatic nodes.

CONFLICT OF INTERESTS

The authors declare that they have no conflict of interests.

ORCID

Annamaria Agnes http://orcid.org/0000-0003-3814-5726
Alberto Biondi http://orcid.org/0000-0002-2470-7858
Ferdinando M Cananzi http://orcid.org/0000-0002-8227-3373

REFERENCES

1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin*. 2015;65:87-108. https://doi.org/10.3322/caac.21262
2. Siewert JR, Böttcher K, Stein HJ, Roder JD. Relevant prognostic factors in gastric cancer: ten-year results of the German Gastric Cancer Study. *Ann Surg*. 1998;228:449-461. https://doi.org/10.1097/00000658-199810000-00002
3. Wu CW, Hsieh MC, Lo SS, Tsay SH, Lui WY, P'eng FK. Relation of number of positive lymph nodes to the prognosis of patients with primary gastric adenocarcinoma. *Gut*. 1996;38:525-527. https://doi.org/10.1136/gut.38.4.525
4. Kim SG, Seo HS, Lee HH, Song KY, Park CH. Comparison of the differences in survival rates between the 7th and 8th editions of the AJCC TNM staging system for gastric adenocarcinoma: a Single-Institution Study of 5,507 patients in Korea. *J Gastric Cancer*. 2017;17:212-219. https://doi.org/10.5230/jgc.2017.17.e23

5. Sano T, Coit DG, Kim HH, et al. Proposal of a new stage grouping of gastric cancer for TNM classification: International Gastric Cancer Association staging project. *Gastric Cancer*. 2017;20:217-225. https://doi.org/10.1007/s10120-016-0601-9
6. Shu P, Qin J, Shen K, et al. The IGCA staging system is more accurate than AJCC7 system in stratifying survival of patients with gastric cancer in stage III. *BMC Cancer*. 2017;17(1):238. https://doi.org/10.1186/s12885-017-3235-3
7. In H, Solsky I, Palis B, Langdon-Embry M, Ajani J, Sano T. Validation of the 8th edition of the AJCC TNM staging system for gastric cancer using the National Cancer Database. *Ann Surg Oncol*. 2017;24:3683-3691. https://doi.org/10.1245/s10434-017-6078-x
8. Kim C-Y, Yang D-H. Adjustment of N stages of gastric cancer by the ratio between the metastatic and examined lymph nodes. *Ann Surg Oncol*. 2009;16:1868-1874. https://doi.org/10.1245/s10434-009-0430-8
9. Bando E, Yonemura Y, Taniguchi K, Fushida S, Fujimura T, Miwa K. Outcome of ratio of lymph node metastasis in gastric carcinoma. *Ann Surg Oncol*. 2002;9:775-784. https://doi.org/10.1245/ASO.2002.10.011
10. Xiao L-B, Yu J-X, Wu W-H, Xu F-F, Yang S-B. Superiority of metastatic lymph node ratio to the 7th edition UICC N staging in gastric cancer. *World J Gastroenterol*. 2011;17:5123-5130. https://doi.org/10.3748/wjg.v17.i46.5123
11. Xu D, Geng Q, Long Z, et al. Positive lymph node ratio is an independent prognostic factor in gastric cancer after d2 resection regardless of the examined number of lymph nodes. *Ann Surg Oncol*. 2009;16:319-326. https://doi.org/10.1245/s10434-008-0240-4
12. Xu J, Bian Y-H, Jin X, Cao H. Prognostic assessment of different metastatic lymph node staging methods for gastric cancer after D2 resection. *World J Gastroenterol*. 2013;19:1975-1983. https://doi.org/10.3748/wjg.v19.i12.1975
13. Sun Z, Xu Y, Li DM, 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. https://doi.org/10.1002/cncr.24989
14. Spolverato G, Ejaz A, Kim Y, et al. Prognostic performance of different lymph node staging systems after curative intent resection for gastric adenocarcinoma. *Ann Surg*. 2015;262:991-998. https://doi.org/10.1097/SLA.0000000000001040
15. Jian-Hui C, Shi-Rong C, Hui W, et al. Prognostic value of three different lymph node staging systems in the survival of patients with gastric cancer following D2 lymphadenectomy. *Tumour Biol*. 2016;37:11105-11113. https://doi.org/10.1007/s13277-015-4191-7
16. Lee JW, Ali B, Park CH, Song KY. Different lymph node staging systems in patients with gastric cancer from Korean: What is the best prognostic assessment tool? *Medicine (Baltimore)*. 2016;95:e3860. https://doi.org/10.1097/MD.0000000000003860
17. Liu H, Deng J, Zhang R, Hao X, Jiao X, Liang H. The RML of lymph node metastasis was superior to the LODDS forevaluating the prognosis of gastric cancer. *Int J Surg*. 2013;11(5):419-424. https://doi.org/10.1016/j.ijsu.2013.03.009
18. Qiu M, Qiu H, Wang Z, et al. The tumor-log odds of positive lymph nodes-metastasis staging system, a promising new staging system for gastric cancer after D2 resection in China. *PLoS One*. 2012;7:e31736. https://doi.org/10.1371/journal.pone.0031736
19. 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. https://doi.org/10.1245/s10434-013-3356-0
20. Zhao LY, Li CC, Jia LY, et al. Superiority of lymph node ratio-based staging system for prognostic prediction in 2575 patients with gastric cancer: validation analysis in a large single center. *Oncotarget*. 2016;7(32):51069-51081. https://doi.org/10.2139/ssrn.2311310
21. Lee JH, Kang JW, Nam BH, et al. Correlation between lymph node count and survival and a reappraisal of lymph node ratio as a predictor of survival in gastric cancer: A multi-institutional cohort study. *Eur J Surg Oncol*. 2017;43(2):432-439. https://doi.org/10.1016/j.ejso.2016.09.007
22. Wang W, Xu DZ, Li YF, et al. Tumor-ratio-metastasis staging system as an alternative to the 7th edition UICC TNM system in gastric cancer after D2 resection-results of a singleinstitution study of 1343 Chinese patients. *Ann Oncol*. 2011;22:2049-2056. https://doi.org/10.1093/annonc/mdq716
23. Wang X, Appleby D-H, Zhang X, Gan L, Wang J-J, Wan F. Comparison of three lymph node staging schemes for predicting outcome in patients with gastric cancer. *Br J Surg*. 2013;100:505-514. https://doi.org/10.1002/bjs.9014
24. Amin M, Edge SB, Greene FL, eds. *AJCC cancer staging manual*. Springer International Publishing
25. Metze K. The association between overall survival and the total number of dissected lymph nodes: an artifact caused by the surgical pathologist? *Ann Surg*. 2009;249:693-694. https://doi.org/10.1097/01.sla.0000348714.87392.94
26. Liu X, Meltzer SJ. Gastric cancer in the era of precision. *Cell Mol Gastroenterol Hepatol*. 2017;3:348-358. https://doi.org/10.1016/j.jcmgh.2017.02.003
27. Jiang Y, Zhang Q, Hu Y, et al. ImmunoScore signature: a prognostic and predictive tool in gastric cancer. *Ann Surg*. 2018;267(3):504-513. https://doi.org/10.1097/SLA.0000000000002116
28. Jiang X, Hiki N, Nunobe S, et al. Prognostic importance of the inflammation-based Glasgow prognostic score in patients with gastric cancer. *Br J Cancer*. 2012;107(2):275-279. https://doi.org/10.1038/bjc.2012.262

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Agnes A; Biondi A; Cananzi FM, et al. Ratio-based staging systems are better than the 7th and 8th editions of the TNM in stratifying the prognosis of gastric cancer patients: A multicenter retrospective study. *J Surg Oncol*. 2019;119:948–957. https://doi.org/10.1002/jso.25411