ESMO GOOD SCIENCE BETTER MEDICINE BEST PRACTICE

ANNALS OF ONCOLOGY driving innovation in oncology

ORIGINAL ARTICLE

# Deep learning radiomic nomogram can predict the number of lymph node metastasis in locally advanced gastric cancer: an international multicenter study

D. Dong[1,2†], M.-J. Fang[1,2†], L. Tang[3†], X.-H. Shan[4†], J.-B. Gao[5†], F. Giganti[6,7,8†], R.-P. Wang[9], X. Chen[10,11], X.-X. Wang[4], D. Palumbo[8,12], J. Fu[3], W.-C. Li[9], J. Li[5], L.-Z. Zhong[1,2], F. De Cobelli[8,12], J.-F. Ji[13*], Z.-Y. Liu[10*] & J. Tian[1,14,15*]

[1]CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing; [2]School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing; [3]Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Radiology Department, Peking University Cancer Hospital & Institute, Beijing; [4]Department of Radiology, Affiliated People's Hospital of Jiangsu University, Zhenjiang, Jiangsu; [5]Department of Radiology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, China; [6]Department of Radiology, University College London Hospital NHS Foundation Trust, London; [7]Division of Surgery and Interventional Science, Faculty of Medical Sciences, University College London, London, UK; [8]Department of Radiology, Experimental Imaging Centre, San Raffaele Scientific Institute, Milan, Italy; [9]Department of Radiology, Guizhou Provincial People's Hospital, Guiyang, Guizhou; [10]Department of Radiology, Guangdong Provincial People's Hospital/Guangdong Academy of Medical Sciences, Guangzhou, Guangdong; [11]Department of Radiology, Guangzhou First People's Hospital, Guangzhou, Guangdong, China; [12]Vita-Salute San Raffaele University, Milan, Italy; [13]Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Gastrointestinal Cancer Center, Peking University Cancer Hospital & Institute, Beijing; [14]Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, School of Medicine, Beihang University, Beijing; [15]Engineering Research Center of Molecular and Neuro Imaging of Ministry of Education, School of Life Science and Technology, Xidian University, Xi'an, Shaanxi, China

Available online 15 April 2020

**Background:** Preoperative evaluation of the number of lymph node metastasis (LNM) is the basis of individual treatment of locally advanced gastric cancer (LAGC). However, the routinely used preoperative determination method is not accurate enough.

**Patients and methods:** We enrolled 730 LAGC patients from five centers in China and one center in Italy, and divided them into one primary cohort, three external validation cohorts, and one international validation cohort. A deep learning radiomic nomogram (DLRN) was built based on the images from multiphase computed tomography (CT) for preoperatively determining the number of LNM in LAGC. We comprehensively tested the DLRN and compared it with three state-of-the-art methods. Moreover, we investigated the value of the DLRN in survival analysis.

**Results:** The DLRN showed good discrimination of the number of LNM on all cohorts [overall C-indexes (95% confidence interval): 0.821 (0.785—0.858) in the primary cohort, 0.797 (0.771—0.823) in the external validation cohorts, and 0.822 (0.756—0.887) in the international validation cohort]. The nomogram performed significantly better than the routinely used clinical N stages, tumor size, and clinical model ($P < 0.05$). Besides, DLRN was significantly associated with the overall survival of LAGC patients ($n = 271$).

**Conclusion:** A deep learning-based radiomic nomogram had good predictive value for LNM in LAGC. In staging-oriented treatment of gastric cancer, this preoperative nomogram could provide baseline information for individual treatment of LAGC.

**Key words:** deep learning, locally advanced gastric cancer, lymph node metastasis, radiomic nomogram

## INTRODUCTION

Gastric cancer is the third leading cause of death from cancer worldwide.[1] The gastric cancer incidence rates in Asia, Eastern Europe, and South America are relatively high.[2,3]

Locally advanced gastric cancer (LAGC), characterized by wall invasion deeper than the submucosa, is associated with a high rate of lymph node metastasis (LNM) and poor clinical outcomes.[4] According to the 8th American Joint Committee on Cancer (AJCC) TNM staging system, the severity of lymph node (LN) involvement is classified based on the number of LNMs as N0 (no LNM), N1 (1—2 LNMs), N2 (3—6 LNMs), N3a (7—15 LNMs), and N3b (>15 LNMs).[5]

*Correspondence to: Prof. Jie Tian, CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, No. 95 Zhongguancun East Road, Hai Dian District, Beijing 100190, China. Tel: +86-10-82618465
E-mail: jie.tian@ia.ac.cn (J. Tian).
*Prof. Zai-Yi Liu, Guangdong Provincial People's Hospital/Guangdong Academy of Medical Sciences, No. 106 Zhongshan Er Road, Guangzhou 510080, China. Tel: +86-20-83870125
E-mail: zyliu@163.com (Z.-Y. Liu).
*Prof. Jia-Fu Ji, Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Gastrointestinal Cancer Center, Peking University Cancer Hospital & Institute, No. 52 Fu Cheng Road, Hai Dian District, Beijing 100142, China. Tel: +86-10-88196598
E-mail: jijiafu@hsc.pku.edu.cn (J.-F. Ji).

†These authors contributed equally to this work.

Accurate preoperative N staging is one of the bases of individual treatment of LAGC. Patients with different N stages have significantly different prognosis and may need a different extent of lymphadenectomy or neoadjuvant treatment.[4] The European prospective randomized Dutch trial showed that extended lymphadenectomy (D2) had a superior survival than limited lymphadenectomy (D1) in LAGC patients with N2 stage.[6] The European Society for Medical Oncology (ESMO) and National Comprehensive Cancer Network (NCCN) guidelines recommend preoperative N staging using medical imaging.[3,4] In particular, computed tomography (CT) imaging has been routinely used for preoperative N staging, with enlarged and round-shaped LNs as a sign of LNM.[5] However, the accuracy of CT is ~50%–70% for LNM,[7] which is unsatisfactory.

Radiomics is an emerging technique that converts standard-of-care medical images into hand-crafted radiomic features and then selects critical features as a signature for quantitative cancer diagnostics.[8–11] Radiomic nomogram, a graphic representation of model that combines radiomic signature and clinical characteristics, has improved the prediction ability of peritoneal metastasis in LAGC.[12] In combination with deep learning features automatically learned from convolutional neural networks, radiomics showed excellent performance in cancer prognosis.[13] However, the use of deep learning radiomics to predict N stages in LAGC has yet to be reported.

To address this, we aimed to develop a deep learning radiomic nomogram (DLRN) for N staging in LAGC. We focused on preoperatively discriminating pathologic N0, N1, N2, N3a, and N3b, because an accurate staging is the basis of individual treatment.

## PATIENTS AND METHODS

### Patients

This retrospective study was approved by the Institutional Review Board of all participating hospitals, and the requirement for informed consent was waived.

Patients were enrolled based on the following inclusion criteria: (i) pathologically diagnosed as LAGC (pT2-4aNxM0); (ii) D2 LN dissection with at least 16 LNs during the surgery; (iii) CT carried out less than 2 weeks before surgery. Patients were excluded based on the following criteria: (i) preoperative therapy (radiotherapy, chemotherapy, or other treatments); (ii) previous abdominal malignancies or inflammatory diseases; (iii) difficult to segment the tumor because of unsatisfactory gastric distention; (iv) artifacts on CT images seriously deteriorating the observation of LNs.

As shown in Figure 1 and supplementary A1, available at *Annals of Oncology* online, 679 LAGC patients were enrolled from five centers in China and divided into four cohorts: a primary cohort for training (PC, $n = 225$) and three validation cohorts (VC1, $n = 178$; VC2, $n = 145$; VC3, $n = 131$). An international validation cohort (IVC, $n = 51$) was collected from Italy. Besides, a follow-up cohort ($n = 271$) was used for survival analysis in LAGC.

### Clinical characteristics

The clinical characteristics of the patients are presented in supplementary Table S1, available at *Annals of Oncology* online. The gold standard for N stages was pathologically assessed after surgery. The clinical N and clinical T stages were determined based on preoperative CT images by experienced radiologists, according to the 8th AJCC TNM staging system.[5,14]

### CT imaging

All patients in PC, VC1, and VC2 underwent both unenhanced and biphasic (arterial and venous phase) contrast-enhanced CT before surgery. Patients in VC3 and IVC underwent only biphasic contrast-enhanced CT. The CT image-acquisition settings are presented in supplementary A2 and Table S2, available at *Annals of Oncology* online.

### Procedures

Figure 1 shows the flowchart of this study. The DLRN modeling pipeline is shown in Figure 2.

### Tumor region segmentation

Tumor regions of interest (ROIs) were manually delineated on multiphase CT images by an experienced radiologist (reader 1). For each CT phase, only one slice with the largest tumor area was chosen visually by the radiologist and a two-dimensional ROI of the tumor was delineated using ITK-SNAP software (version 3.6.0; http://www.itksnap.org). After 3 months, 30 patients in the PC were randomly selected, and their ROIs were segmented again by reader 1 and another radiologist (reader 2) to construct two resegmentation datasets for the assessment of intrareader/interreader reproducibility of radiomic features.

### Radiomic feature extraction

A total of 112 deep learning features and 289 hand-crafted features were extracted from each ROI, totaling 1203 features from the three ROIs per patient (supplementary A3, available at *Annals of Oncology* online). We adapted the DenseNet-201 architecture to develop our deep convolutional neural networks (DCNNs) for deep learning feature extraction.[15] The hand-crafted features included shape, global texture, and local texture.

### Radiomic signature building

Feature selection and signature building were carried out in PC (supplementary A4, available at *Annals of Oncology* online). Three signatures were respectively built from the three ROIs as follows: (i) intraclass/interclass correlation coefficients (ICCs) and coefficient of variation (CV) were calculated on the resegmentation dataset and a simulated slice thickness dataset (supplementary A5, available at *Annals of Oncology* online), respectively. The stable features with ICCs $> 0.8$ and CV$< 15\%$ were selected to adapt different segmentations and different slice thicknesses. (ii) The features were divided into several clusters by

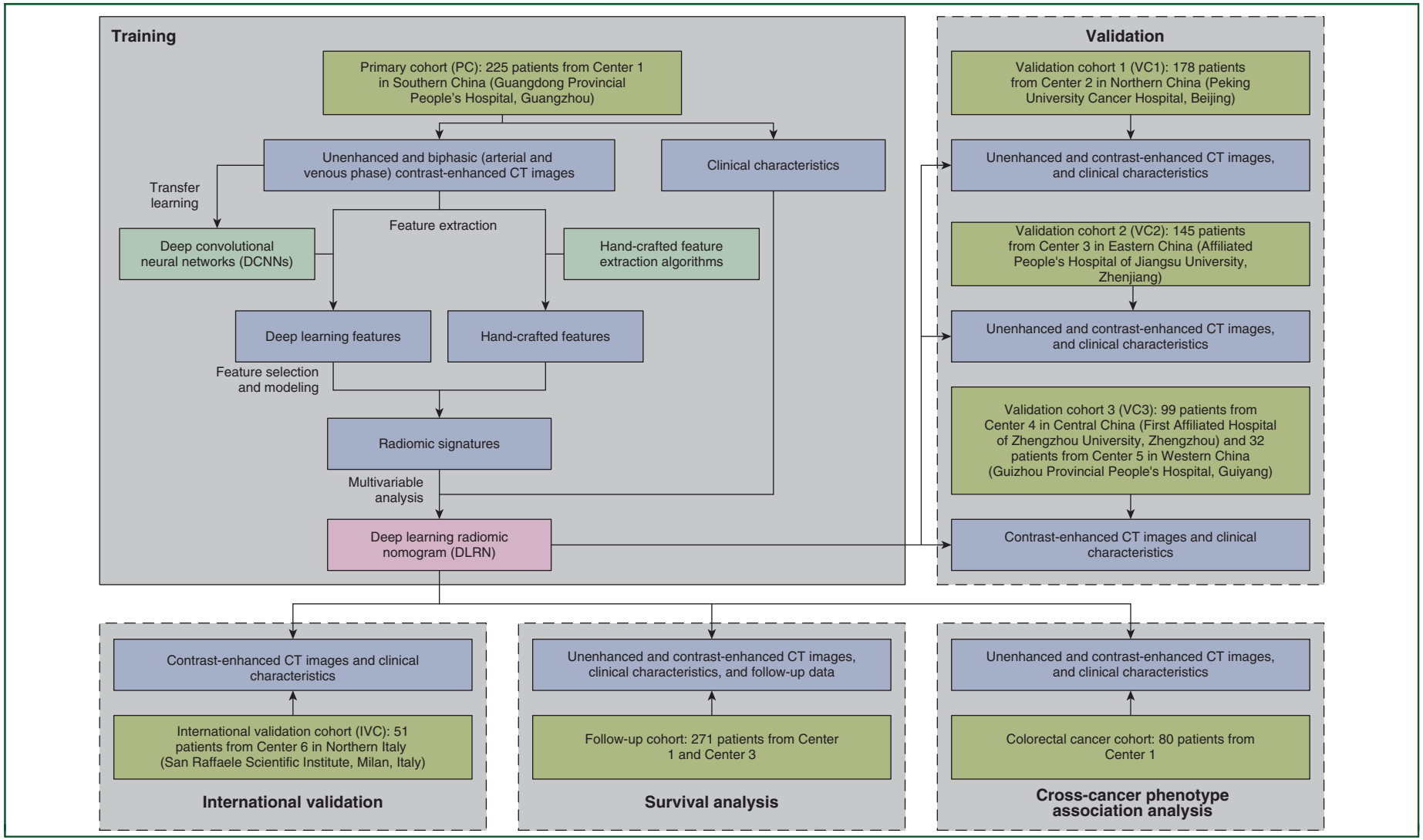


**Figure 1. Flowchart of this international multicenter study.**
CT, computed tomography.

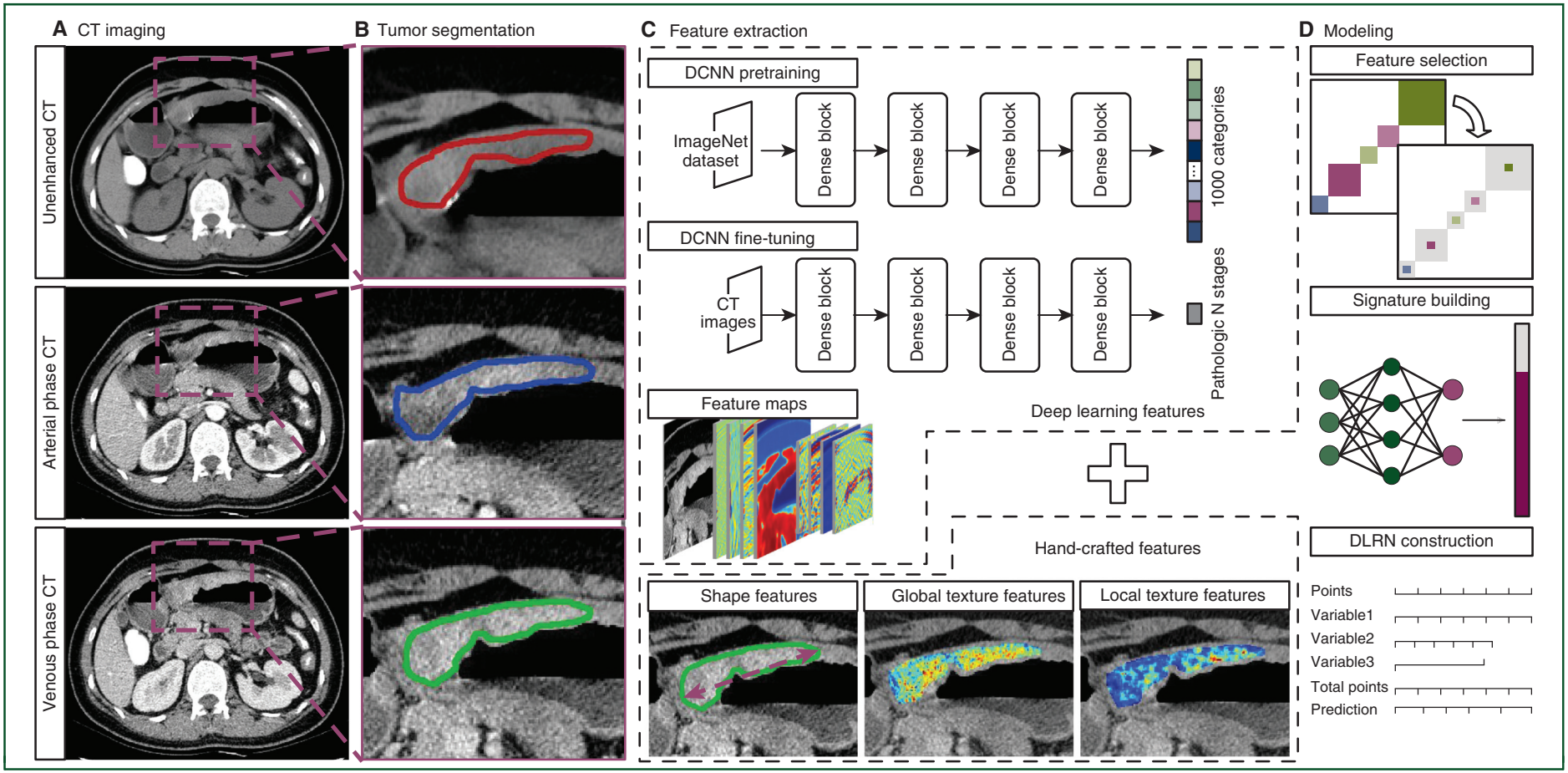


**Figure 2. Workflow of deep learning radiomic nomogram (DLRN) modeling for N staging in locally advanced gastric cancer (LAGC) patients.**
CT, computed tomography; DCNN, deep convolutional neural network.


https://doi.org/10.1016/j.annonc.2020.04.003

hierarchical clustering and the most representative medoid feature in each cluster was reserved. (iii) Three methods, including support vector machine, artificial neural network, and random forest, were compared and the best method was used to construct three predictive signatures.

### DLRN construction

Univariate analysis was used to select statistically significant clinical characteristics ($P < 0.05$). Multivariable linear regression analysis was conducted to build the DLRN from the clinical characteristics and radiomic signatures. We mainly considered contrast-enhanced radiomic signatures in our DLRN; however, the incremental predictive value of unenhanced radiomic signature to DLRN was also investigated using the net reclassification index (NRI).

The association of DLRN score with pathologic N stages was assessed using Spearman's correlation analysis. Logistic regression was used to predict the probability belonging to each N stage with the DLRN score. In addition, a classification procedure was proposed based on cut-offs of the logistic regressions above to split patients into subgroups of N stages. Furthermore, multivariable logistic regression was carried out to build a clinical model based on clinical characteristics for comparison.

### Performance evaluation

Harrell's C-indexes[16] of the DLRN, radiomic signatures, significant clinical characteristics, and clinical model were compared in all cohorts. The confusion matrix of DLRN was also depicted. Moreover, stratification analysis was presented on clinical characteristics and CT scan parameters.

Furthermore, we carried out subgroup analysis and calculated pairwise C-indexes for discriminating non-N0 versus N0, N2—3b versus N0—1, N3a—3b versus N0—2, and N3b versus N0—3a. The calibration curve was plotted to assess the calibration of the DLRN on the subgroup analysis. Among the subgroup analysis, non-N0 versus N0 is of special concern because it may determine the surgical strategy for lymphadenectomy. Decision curve analysis was conducted to evaluate the clinical usefulness of our DLRN in guiding lymphadenectomy by quantifying the net benefits.

We further validated our DLRN on the Italian cohort using Spearman's correlation coefficient and overall C-index. Besides, we evaluated the association between DLRN score and overall survival (OS) in the follow-up LAGC cohort using Kaplan—Meier curves.

### Statistical analysis

Statistical analysis was conducted with R software (version 3.5.0; http://www.Rproject.org; R Foundation for Statistical Computing, Vienna, Austria) and MATLAB. A two-sided $P < 0.05$ was used as the criterion of statistically significant difference. In the univariate analysis, the differences in clinical characteristics between the patients in different groups were assessed using independent *t* test or Mann—Whitney *U* test for continuous variables and Fisher's exact test or chi-square test for categorical variables. Analysis of variance and Kruskal—Wallis H test were implemented for comparing three or more groups.

## RESULTS

Chi-square test and *t* test showed that there was no significant difference in sex or age between PC and VCs in China ($P > 0.05$) except for patients' age in VC2 ($P = 0.0040$). As shown in supplementary Table S1, available at *Annals of Oncology* online, the pathologic N stage was significantly associated with tumor size, clinical N stage, clinical T stage, sex, and CA19-9 in the PC ($P < 0.05$).

During the radiomic signature building step (supplementary A6 and Figure S1, available at *Annals of Oncology* online), support vector machine was optimally selected to build three radiomic signatures, including arterial signature (six features), venous signature (six features), and unenhanced signature (seven features). The final features are shown in supplementary Table S3 and supplementary A6, available at *Annals of Oncology* online.

The multivariable linear regression analysis in the PC showed that arterial signature, venous signature, and clinical N stage were independent predictors for pathologic N stage (supplementary Table S4, available at *Annals of Oncology* online), while the clinical T stage, tumor size, sex, and CA19-9 were not significant. These predictors were combined into the DLRN (Figure 3A). The NRI analysis revealed that the addition of an unenhanced signature into DLRN did not show significantly better performance (NRI 0.0482; $P = 0.1870$).

As shown in Figure 3B, there was a significant positive correlation between DLRN score and pathologic N stage, which was also confirmed by Spearman's correlation coefficients (0.626—0.718, $P < 0.0001$) in supplementary Table S5, available at *Annals of Oncology* online, and the confusion matrixes in supplementary Figure S2, available at *Annals of Oncology* online. As shown in Table 1, the DLRN showed a good discrimination of N stages in PC [overall C-index 0.821, 95% confidence interval (CI) 0.785—0.858], VC1 (0.777, 0.735—0.819), VC2 (0.817, 0.775—0.860), and VC3 (0.787, 0.737—0.838). Moreover, the DLRN performed significantly better than the clinical N stage, tumor size, and the clinical model (supplementary Table S4, available at *Annals of Oncology* online) on all the external VCs in China with $P < 0.05$. The stratification analysis showed that the performance of our DLRN was not affected by the age, sex, Lauren type, tumor location, the version of CT system, and slice thickness (supplementary Figure S3 and supplementary A7, available at *Annals of Oncology* online).

The calibration curves of the subgroup analysis showed good agreement between the DLRN predicted outcomes and the real N stages (Figure 3C). Moreover, the DLRN could well discriminate non-N0 from N0 groups in all cohorts (C-indexes: 0.777—0.821, supplementary Table S5, available at *Annals of Oncology* online). If we use this model to guide lymphadenectomy (non-N0 patients receive lymphadenectomy and N0 patients do not), as shown in Figure 3D, the decision curves indicated that the DLRN could add more benefit to patients

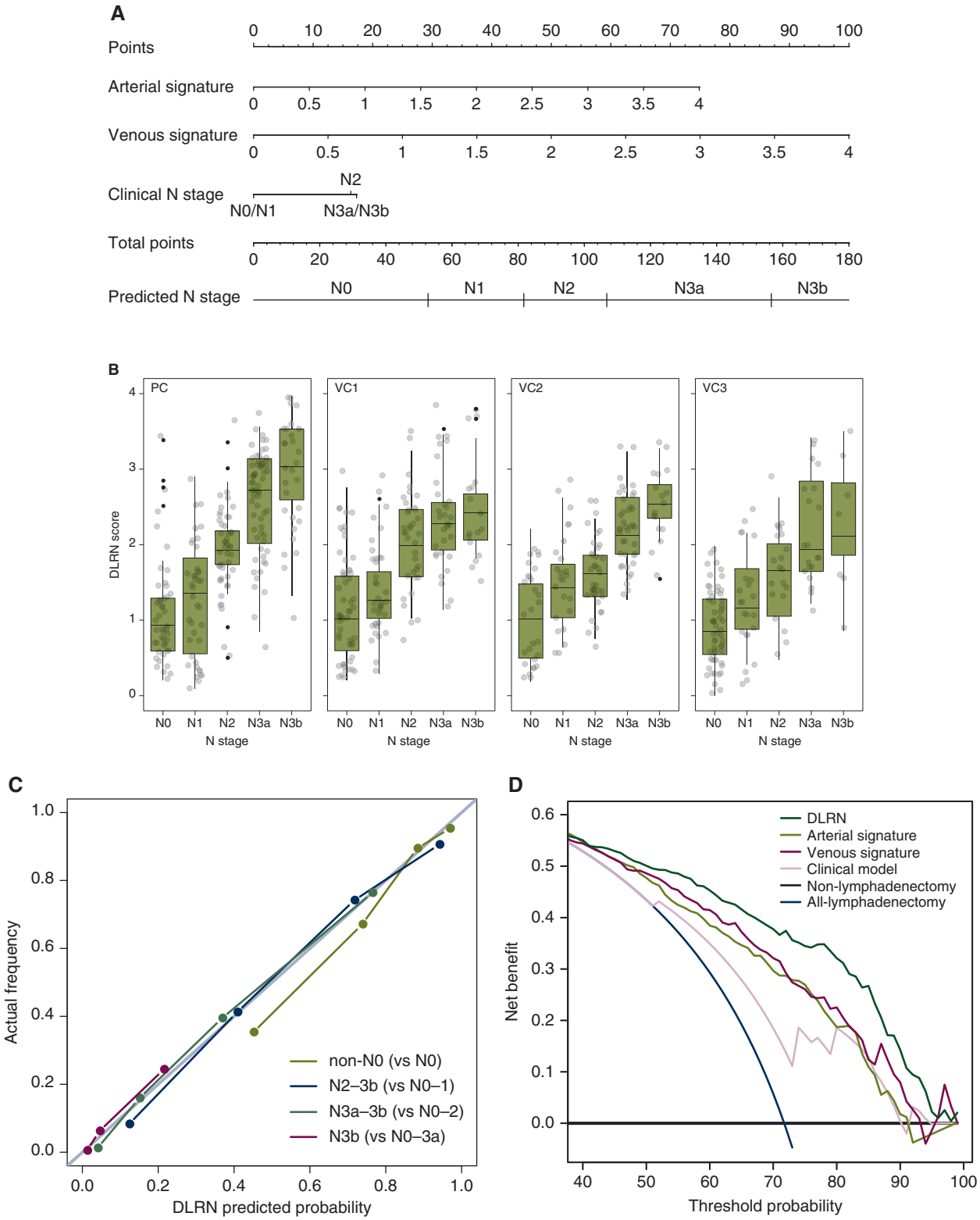


**Figure 3. Deep learning radiomic nomogram (DLRN) and its performance.**
(A) DLRN with two contrast-enhanced radiomic signatures and clinical N stage. The points of arterial signature, venous signature, and clinical N stage are obtained based on the top 'points' bar with scale of 0—100. Then, the total point is calculated by summing the three points. The predicted N stage is obtained by mapping the total point to the 'total points' bar and the 'predicted N stage' bar. (B) Box plots showing patterns of correlation between pathologic N stages and DLRN in PC, VC1, VC2, and VC3. (C) Calibration curves of DLRN in subgroup analysis on discriminating non-N0 versus N0, N2—3b versus N0—1, N3a—3b versus N0—2, and N3b versus N0—3a. (D) Decision curve analysis for guiding lymphadenectomy using DLRN, arterial signature, venous signature, clinical model, non-lymphadenectomy scheme, and all-lymphadenectomy scheme.


https://doi.org/10.1016/j.annonc.2020.04.003

**Table 1. Overall C-index of DLRN and other predictors**

| | PC | VC1 | VC2 | VC3 | All VCs |
|---|---|---|---|---|---|
| DLRN | 0.821 (0.785–0.858) | 0.777 (0.735–0.819) | 0.817 (0.775–0.860) | 0.787 (0.737–0.838) | 0.797 (0.771–0.823) |
| Arterial signature | 0.766 (0.719–0.812) | 0.738 (0.688–0.787) | 0.761 (0.704–0.818) | 0.716 (0.652–0.780) | 0.738 (0.705–0.770) |
| Venous signature | 0.785 (0.744–0.826) | 0.719 (0.667–0.770) | 0.732 (0.676–0.789) | 0.739 (0.678–0.799) | 0.739 (0.708–0.770) |
| Unenhanced signature | 0.782 (0.740–0.824) | 0.729 (0.681–0.776) | 0.676 (0.611–0.740) | | 0.697 (0.657–0.736)[a] |
| Clinical N stage | 0.679 (0.629–0.730) | 0.685 (0.629–0.741) | 0.698 (0.631–0.766) | 0.709 (0.619–0.800) | 0.705 (0.669–0.742) |
| Clinical model | 0.689 (0.642–0.736) | 0.652 (0.593–0.711) | 0.671 (0.605–0.737) | 0.732 (0.661–0.804) | 0.675 (0.638–0.713) |
| Tumor size | 0.666 (0.619–0.714) | 0.666 (0.610–0.722) | 0.673 (0.616–0.730) | 0.638 (0.565–0.711) | 0.664 (0.630–0.699) |

DLRN, deep learning radiomic nomogram.
[a] The value was calculated based on VC1 and VC2.

than single signatures, clinical model, non-lymphadenectomy scheme, and all-lymphadenectomy scheme.

Clinicians may be interested in how many patients with CT-diagnosed N0 disease will be upstaged with DLRN (non-N0 by pathology). These cases could be named as occult LNM, which are with no typical CT signs (i.e. enlarged LN). The experimental results showed that DLRN could well detect these patients with occult LNM [81.7% (76/93) upgraded].

We further validated our DLRN on the non-Asian cohort IVC (supplementary Table S6, available at *Annals of Oncology* online). The DLRN also showed good discrimination of N stage in IVC (overall C-index 0.822, 95% CI 0.756–0.887).

We evaluated the prognostic value of DLRN in the follow-up LAGC cohort (supplementary Table S6, available at *Annals of Oncology* online). The DLRN yielded a predictive accuracy for OS (C-index 0.646, 95% CI 0.596–0.696, $P < 0.0001$). Patients with high DLRN score displayed worse OS [per 1 increase; hazard ratio (HR) 1.982, 95% CI 1.592–2.467, $P < 0.0001$]. As shown in Figure 4, the Kaplan–Meier curves divided by the median value of DLRN score were significantly different (log-rank test $P < 0.0001$). Further, we carried out univariate analysis and multivariate Cox regression analyses on DLRN and clinical characteristics. As shown in supplementary Table S7, available at *Annals of Oncology* online, the DLRN had the highest C-index of predicting OS in the univariate analysis. The Cox regression identified the DLRN and invaded site as the independent prognostic factors. The final Cox regression model yielded a C-index of 0.656 (95% CI 0.606–0.705).

## DISCUSSION

This study was an international multicenter collaboration aimed at predicting the number of LNM in LAGC. Our DLRN showed high predictive ability and reproducibility across different centers. Moreover, the ROI segmentation and DLRN score calculation require less than 5 additional minutes per patient during a normal reporting session, which makes DLRN an easy-to-use tool for clinicians. We have already uploaded the model and several examples of CT images on our website (http://www.radiomics.net.cn/

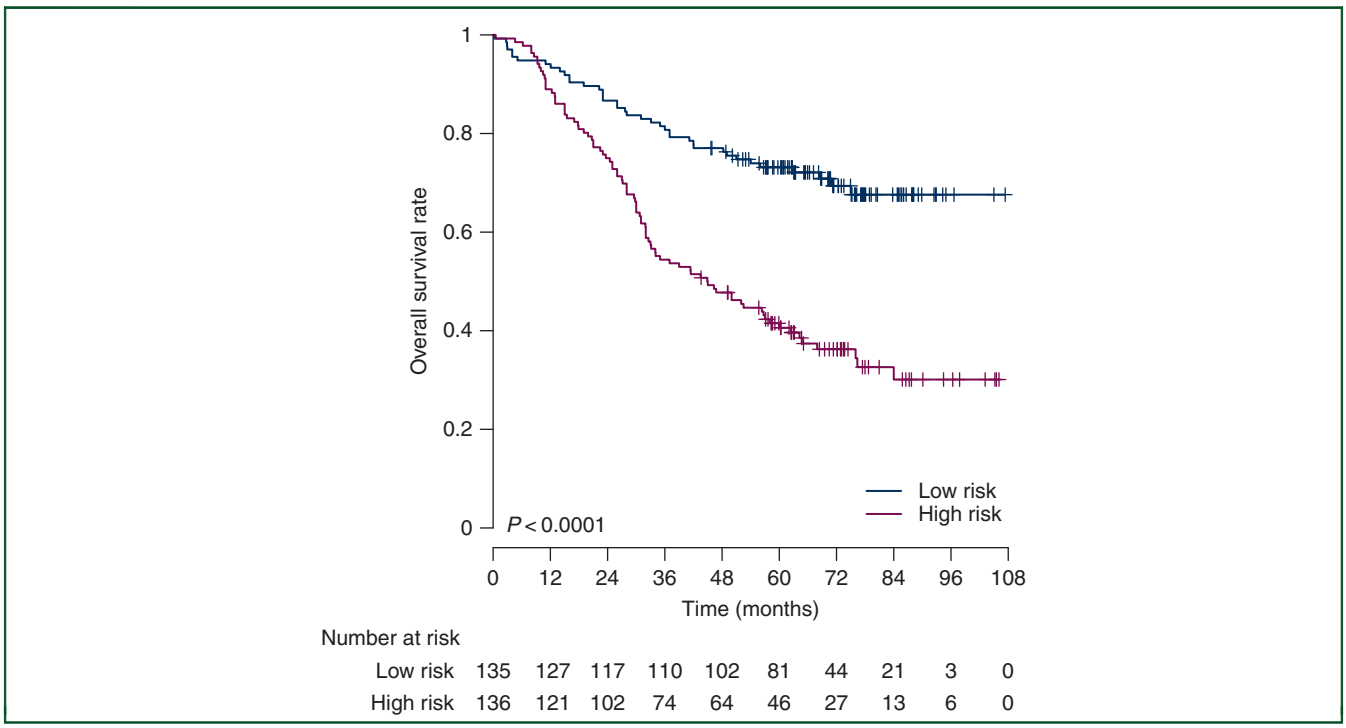


**Figure 4. Kaplan–Meier survival curves of overall survival on the follow-up locally advanced gastric cancer cohort.**

platform.html) as well as on Zenodo (https://doi.org/10.5281/zenodo.3701430) for open access.

According to the latest TNM staging system, regional LNs with enlarged short-axis diameter ≥1 cm and other abnormal signs on imaging are suspicious for nodal involvement.[5] However, this standard (the clinical N stage in this study) showed relatively poor performance in our cohorts. By contrast, our DLRN performed significantly better than the routinely used clinical N stage. Moreover, 81.7% of occult LNMs with no typical CT signs (missed by the radiologists) were detected by DLRN, which indicated that our model could be a supplement to current staging scheme.

Our DLRN may help tailor neoadjuvant therapy, lymphadenectomy, or extent of lymphadenectomy in LAGC. There is growing interest in the use of neoadjuvant therapy before surgery for LAGC.[4] Patients with neoadjuvant chemotherapy were proven to have fewer LNMs after surgery than those without.[17] This finding suggests that our preoperative DLRN may be helpful for tailoring neoadjuvant regimens. Moreover, the decision curve analysis showed that using our DLRN to guide lymphadenectomy could provide more benefit to patients than both non-lymphadenectomy and all-lymphadenectomy schemes. Even for patients with lymphadenectomy, there is a long-running debate over which lymphadenectomy extent (D0, D1, or D2) could be beneficial for patients.[4,6,18] Our nomogram is able to evaluate the number of LNM preoperatively, which could, in turn, assist in choosing the extent of lymphadenectomy.

This study was carried out on three tumor ROIs from multiphase CT images rather than on one ROI from a single CT phase. Although the three ROIs showed different shapes and contents, their radiomic signatures were all significantly associated with the pathologic N stages ($P < 0.001$). Furthermore, 13 of 19 selected features in the three signatures were deep learning features, indicating that the DCNNs could extract correlative quantitative representation reflecting the extent of LNM. As shown in supplementary Figure S4, available at *Annals of Oncology* online, the activation maps of DCNN could highlight some regions of the tumors with a large number of LNM, while the same region was suppressed in tumors with small number of LNM. We suspect that the highlighted regions in the activation maps may be relevant to cancer progression. Besides, the global texture features were also adopted in the radiomic signatures, which might reflect the heterogeneity and invasiveness of the tumor. For example, the 'gray-level co-occurrence matrix (GLCM) dissimilarity' feature qualifies the global distribution characteristics of gray-level variability in tumor ROI. The feature 'GLCM cluster_tendency' tends to emphasize the ROI with significant textural patterns.

Another finding was that our DLRN was significantly associated with the OS of LAGC patients. Previous studies have proven that LAGC patients with different N stages had different prognosis.[4] Our results further validated the association with N stages as well as the prognosis value of our DLRN. Furthermore, we conducted a cross-cancer analysis and transferred our model to a colorectal cancer cohort ($n = 80$). Interestingly, all three radiomic signatures had the potential to discriminate LNM of colorectal cancer (supplementary A8 and supplementary Figure S5, available at *Annals of Oncology* online), indicating that other gastrointestinal cancers might have similar phenotypes with LNM.

Our study has some limitations. First, this study involved a large number of patients from China but a small number of patients from Italy. A further prospective study on other Asian and large-scale non-Asian populations should be investigated. Second, gastric cancer can have different etiology and biology in different countries or races; how this influences our nomogram remains unclear. However, mixing patients from different countries/races for training may improve the performance of the model. Third, besides CT, endoscopic ultrasonography is also recommended for N staging.[4] The combination of endoscopic ultrasonography and CT may improve N staging accuracy. Fourth, the 2D features in one single slice rather than 3D features were used. Although the operation is more convenient for the radiologist, the 2D segmentation may not be representative of the entire tumor and some features may be affected from 2D versus 3D. Finally, gastric cancer with microsatellite instability (~10% percentage) is less likely to have LNM but is likely to have enlarged LNs due to immune infiltrate[19] and its contribution in the nomogram should be further investigated.

In conclusion, a DLRN had good predictive ability for N staging in LAGC, which could provide basic information for individual diagnosis and treatment in LAGC.

## ACKNOWLEDGEMENTS

We thank Dr Lambin Philippe and Dr Henry C Woodruff from the D-Lab, GROW Research Institute for Oncology, Maastricht University, The Netherlands, for their careful review of the paper.

## FUNDING

This work was supported by the National Natural Science Foundation of China [grant numbers 91959130, 81971776, 81771924, 81930053, 81771912, 81671682, 91959205, 81601469, 81701687, 81227901], The National Science Fund for Distinguished Young Scholars [grant number 81925023], National Key R&D Program of China [grant numbers 2017YFA0205200, 2017YFC1308700, 2018YFC0910700, 2017YFC1309100, 2017YFC1309101, 2017YFC1309104], the Beijing Natural Science Foundation [grant numbers L182061, Z180001], and the Youth Innovation Promotion Association CAS [grant number 2017175]. FG is funded by the UCL Graduate Research Scholarship and the Brahm PhD scholarship in memory of Chris Adams (no grant number).

## DISCLOSURE

The authors have declared no conflicts of interest.

## REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68:394–424.

2. Shen L, Shan Y, Hu H, et al. Management of gastric cancer in Asia: resource-stratified guidelines. *Lancet Oncol*. 2013;14:e535—e547.
3. Smyth EC, Verheij M, Allum W, et al. Gastric cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2016;27:v38—v49.
4. National Comprehensive Cancer Network (NCCN) guidelines. Available at http://www.nccn.org/. Accessed January 10, 2018.
5. Amin MB, Edge SB, Greene F, et al., eds. *AJCC Cancer Staging Manual*. Basel, Switzerland: Springer; 2017.
6. Schwarz RE, Smith DD. Extended lymph node dissection for gastric cancer: who may benefit? Final results of the randomized Dutch gastric cancer group trial. *J Clin Oncol*. 2004;22:2069—2077.
7. Kim HJ, Kim AY, Oh ST, et al. Gastric cancer staging at multi-detector row CT gastrography: comparison of transverse and volumetric CT scanning. *Radiology*. 2005;236:879—885.
8. Lambin P, Riosvelazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. 2012;48:441—446.
9. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol*. 2017;14:749—762.
10. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun*. 2014;5:4006.
11. Huang Y, Liang CC, He L, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. *J Clin Oncol*. 2016;34:2157—2164.
12. Dong D, Tang L, Li ZY, et al. Development and validation of an individualized nomogram to identify occult peritoneal metastasis in patients with advanced gastric cancer. *Ann Oncol*. 2019;30:431—438.
13. Peng H, Dong D, Fang MJ, et al. Prognostic value of deep learning PET/CT-based radiomics: potential role for future individual induction chemotherapy in advanced nasopharyngeal carcinoma. *Clin Cancer Res*. 2019;25:4271—4279.
14. 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.
15. Huang G, Liu Z, Pleiss G, et al. Convolutional networks with dense connectivity. *IEEE Trans Pattern Anal*. 2019. https://doi.org/10.1109/TPAMI.2019.2918284.
16. Moons KGM, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*. 2015;162:W1—W73.
17. Schuhmacher C, Gretschel S, Lordick F, et al. Neoadjuvant chemotherapy compared with surgery alone for locally advanced cancer of the stomach and cardia: European organisation for research and treatment of cancer randomized trial 40954. *J Clin Oncol*. 2010;28:5210—5218.
18. Degiuli M, Sasako M, Ponti A, et al. Randomized clinical trial comparing survival after D1 or D2 gastrectomy for gastric cancer. *Br J Surg*. 2014;101:23—31.
19. Corso G, Pedrazzani C, Marrelli D, et al. Correlation of microsatellite instability at multiple loci with long-term survival in advanced gastric carcinoma. *Arch Surg*. 2009;144:722—727.