# Risk factors of tumor invasion and node metastasis in early gastric cancer with undifferentiated component: a multicenter retrospective study on biopsy specimens and clinical data

Yi Zou[1], Long Wu[2], Yubin Yang[3], Xin Shen[4], Chunpeng Zhu[5]

[1]Department of Pathology, Second Affiliated Hospital of Zhejiang University, School of Medicine, Hangzhou 310009, China; [2]Department of Pathology, Union Hospital of Fujian Medical University, Fuzhou 350001, China; [3]Department of Pathology, Second Affiliated Hospital of Fujian Medical University, Quanzhou 362000, China; [4]College of Computer Science and Technology, Zhejiang University, Hangzhou 310027, China; [5]Department of Gastroenterology, Second Affiliated Hospital of Zhejiang University, School of Medicine, Hangzhou 310009, China

*Contributions:* (I) Conception and design: C Zhu, Y Zou; (II) Administrative support: C Zhu; (III) Provision of study materials or patients: C Zhu, Y Zou; (IV) Collection and assembly of data: Y Zou, L Wu, Y Yang, X Shen; (V) Data analysis and interpretation: Y Zou, L Wu, Y Yang, X Shen; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

*Correspondence to:* Chunpeng Zhu. Department of Gastroenterology, Second Affiliated Hospital of Zhejiang University, School of Medicine, No. 88 Jiefang Road, Shangcheng District, Hangzhou 310009, China. Email: zhuchunpeng@zju.edu.cn.

**Background:** Early gastric cancer (EGC) with undifferentiated component (UDC) is a more aggressive entity, where the significance of preoperative data to tumor invasion and lymph node metastasis (LNM) remains unclarified.

**Methods:** A total of 5,020 GC patients undergoing radical gastrectomy in three centers were reviewed, of which, EGC with UDC in preoperative biopsy specimens were enrolled. The histology of biopsy and surgical specimens was graded according to the proportion of UDC and signet ring cells (SRCs). Risk factors of tumor invasion and LNM were evaluated with histological, clinical and demographic data.

**Results:** Lower body mass index (BMI), melena and larger tumor size were the independent preoperative risk factors of both LNM and LVI, while ulcerative lesion (UL) and the lower third stomach were only correlated with LNM. No relevance was found between the histological features of biopsy specimens and LNM, but SRC or >50% UDC lowered the risk of lymphovascular invasion (LVI) and/or submucosal (SM) invasion. When surgical data (depth of invasion and LVI included) were added, lower BMI, melena and the lower third stomach were still the independent preoperative risk factors of LNM, and LVI, SRC and SM invasion also showed relevance to LNM. The performance of predictive models using pre- or postoperative histological data was comparable.

**Conclusions:** The preoperative data were significantly relevant to tumor invasion and LNM, showing comparable risk strength with surgical specimens in histology.

**Keywords:** Early gastric cancer (EGC); poorly differentiated adenocarcinoma; signet ring cells; lymph node metastasis (LNM)

## Introduction

Gastric cancer (GC) is the fifth most common cancer and the third leading cause of cancer death worldwide, responsible for over 1,000,000 new cases in 2018 and 1 in every 12 deaths globally (1). Although the high mortality GC causes, the early diagnosis and intervention improve the 5-year survival rate from 20–30% in patients at late stage (2) to 70–90% in early gastric cancer (EGC) (3,4).

**Page 2 of 13**

Zou et al. Preoperative risk factors of early gastric cancer

Recently, endoscopic submucosal dissection (ESD) plays an increasingly important role in EGC therapy. However, in the absolute indication for ESD established by the Japanese Gastric Cancer Association, only non-ulcerated, well-differentiated and mucosal ≤2 cm lesions were included, and the expanded criteria, including non-lymphovascular-infiltrated, larger, ulcerated, and undifferentiated mucosal lesions, or differentiated lesions with slight submucosal (SM) invasion, remain controversial (5). How to identify patients with high risk of lymph node metastasis (LNM) accurately is the core problem so far.

Many scholars have tried to identify the risk factors of LNM in previous studies. It was reported that the rates of LNM in intramucosal or EGC varied greatly, ranging from 0.4% to 24% (2,3,6,7). According to histological manifestation, EGC is divided into differentiated and undifferentiated type, the former covers papillary adenocarcinoma, well and moderately-differentiated tumor and the latter includes poorly-differentiated tumor and signet ring cell carcinoma (SRCC) (8). For total or differentiated EGC, various demographic and clinicopathological features were picked out as the risk factors of LNM from original research, meta-analysis or data of Surveillance, Epidemiology and End Results program (SEER), such as patients' gender (6), age (9), tumor location (3,6), gross pattern (10), tumor size (10), ulcerative lesion (UL) (2), depth of invasion (11), histological differentiation (10), and lymphovascular invasion (LVI) (11). For pure undifferentiated EGC, gender, age, tumor location, tumor size, gross pattern, LVI, and depth of invasion were also identified as the risk factors of LNM in previous studies (12-14). However, evidence is still lacking concerning differentiated EGC accompanied with undifferentiated components (UDC) (5), as the presence of UDC showed significantly different biological behaviors when their proportion varied (2,11,15-17).

More importantly, nearly all the previous studies were based on the clinicopathological features of ESD/surgical specimens, which means all the risk factors mentioned above were postoperative. In this scenario, patients with high risk of LNM could only be identified at the post-ESD stage, and they have to undergo the additional gastrectomy and lymph node dissection after an unnecessary but innocent ESD. Therefore, in the present study, we focused on the relationship between the histological features of biopsy specimens and clinical data, and found the significant association between preoperative data and tumor behaviors.

## Methods

### Patient selection

This was a multi-center retrospective study. The pathological reports of 5,020 patients who underwent radical gastrectomy and lymph node dissection between 2013–2018 in Second Affiliated Hospital of Zhejiang University, Union Hospital of Fujian Medical University and Second Affiliated Hospital of Fujian Medical University were reviewed. Cases met the following inclusion criteria were enrolled in this study: patients were pathologically diagnosed as early gastric adenocarcinoma (pT1 tumor); detailed pathological reports of radical gastrectomy (lymph node dissection included) and hematoxylin & eosin (H&E) slides of matched preoperative biopsy specimens, as well as their clinical data (endoscopic data included) were available; UDC was present in preoperative biopsy specimens (*Figure 1*). The study was approved by the Institutional Review Board of Second Affiliated Hospital of Zhejiang University (2019-ERR-446), Union Hospital of Fujian Medical University (2020KY005) and Second Affiliated Hospital of Fujian Medical University (2019-SAHFMER-208). Patient consent was waived by the institutional review boards, as this study was retrospective and patients' information was protected by blind method.

### Clinicopathological analysis

The histology of biopsy and surgical specimens was graded on the proportion of UDC first. Tumors composed of <50% UDC were classified into U1, those composed of ≥50% but ≤90% UDC were classified into U2, and those composed of >90% UDC were classified into U3. Then, tumors containing SRC were classified into S1, and those without SRC were classified into S0. The information of depth of invasion, LVI and LNM was retrieved from the pathological reports of surgical specimens.

### Demographic and clinical data collection

Age, sex, body mass index [BMI, BMI = weight (kg)/height (m)$^2$], and the presence of weight loss, melena or hematemesis was collected from Electronic Medical Record System. Tumor location (cardia and fundus, body, or, angle and antrum were classified into the upper, middle or lower third stomach, retrospectively), tumor size (the maximal diameter), the presence of UL, and main gross patterns
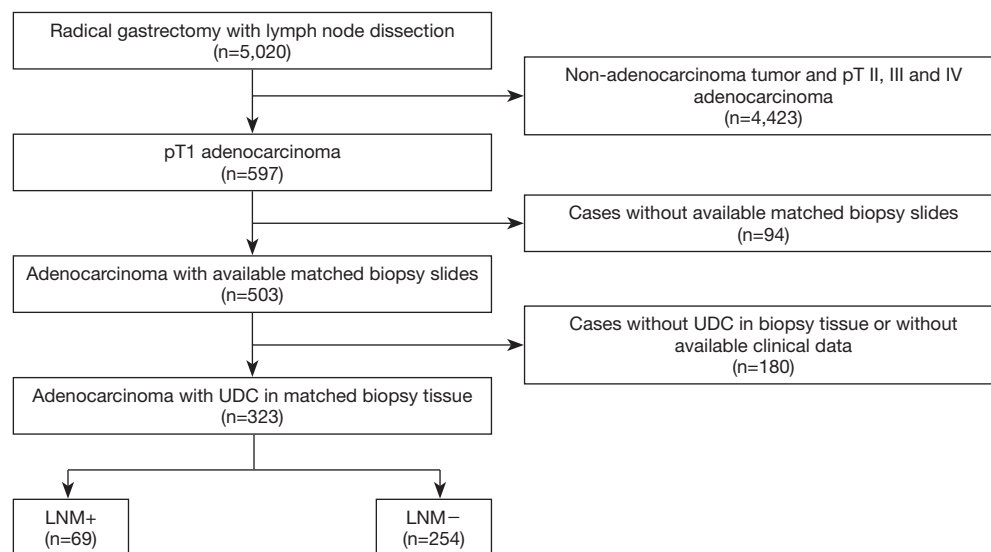
**Figure 1** Flowchart of case selection. +, presence; –, absence. UDC, undifferentiated component; LNM, lymph node metastasis.

evaluated based on the Paris endoscopic classification of superficial neoplastic lesions (18) were retrieved from original endoscopic reports.

### Statistical analysis

The $\chi^2$ test or Fisher's exact test was used to analyze the difference between demographic and clinicopathological ratios, and analysis of variance (ANOVA) was used to compare the quantitative data. Correlation between factors was evaluated by Pearson or Spearman test. Risk analysis was performed using binary logistic regression, and the model performance was evaluated by receiver operating characteristic curve (ROC) curves.

A P value less than 0.05 was considered statistically significant. Statistical analyses were performed using SPSS 22.0 (SPSS Inc., Chicago, IL, USA).

## Results

### Demographic and clinical characteristics

A total of 323 patients were enrolled in this study, of which 247 (76.5%), 42 (13.0%) and 34 (10.5%) were from Second Affiliated hospital of Zhejiang University, Union Hospital and Second Affiliated Hospital of Fujian Medical University, respectively. Patients aged from 24 to 85, with

a slight male predominance. Two hundred and seventeen (67.2%) patients were in normal BMI range (18.5–24 kg/m²) with other 21 (6.5%) and 85 (26.3%) showing low (< 18.5 kg/m²) and high (≥24 kg/m²) BMI. Eighty-six (26.6%) patients presented with at least one of the specific symptoms (weight loss, melena and/or hematemesis). Tumors were relatively small and located at the lower third stomach mostly, followed by the middle and upper third stomach. Most of them showed the gross pattern of IIc, and UL was seen in nearly 40% of the patients (*Table 1*).

LNM was found in 69 (21.4%) patients who presented with melena at a higher rate. Tumors with LNM were located at the lower third stomach more frequently, accompanied with UL and larger size (*Table 1*).

### Clinicopathological characteristics

For the histology of either biopsy or surgical specimens, most cases were classified into grade U3, and nearly half of the cases harbored SRC. No significant histological difference was found between different LNM status. However, tumors with LNM showed deeper invasion and higher LVI rate in surgical specimens. To evaluate the representativeness of biopsy specimens, the correlation of histology between biopsy and corresponding surgical specimens was analyzed, from which significant positive relevance of UDC and SRC grades was found (*Table 2*).

Page 4 of 13

Zou et al. Preoperative risk factors of early gastric cancer

**Table 1** Demographic and clinical characteristics (%)

| Factors | Total | Node metastasis | | P value |
|---|---|---|---|---|
| | | Negative | Positive | |
| Age (year) | | | | |
| Mean ± SD | 56.2±11.6 | 55.8±11.8 | 57.4±10.9 | 0.309 |
| Median | 57 | 56.9 | 57.9 | – |
| Sex | | | | |
| Male | 174 (53.9) | 138 (54.3) | 36 (52.2) | 0.786 |
| Female | 149 (46.1) | 116 (45.7) | 33 (47.8) | |
| BMI | | | | |
| Mean ± SD | 22.5±3.0 | 22.6±3.0 | 22.3±2.9 | 0.562 |
| Median | 22.5 | 22.5 | 22.2 | – |
| Weight loss | | | | |
| – | 261 (80.8) | 210 (82.7) | 51 (73.9) | 0.120 |
| + | 62 (19.2) | 44 (17.3) | 18 (26.1) | |
| Melena | | | | |
| – | 290 (89.8) | 237 (93.3) | 53 (76.8) | *<0.001* |
| + | 33 (10.2) | 17 (6.7) | 16 (23.2) | |
| Hematemesis | | | | |
| – | 312 (96.6) | 247 (97.2) | 65 (94.2) | 0.257 |
| + | 11 (3.4) | 7 (2.8) | 4 (5.8) | |
| Tumor location | | | | |
| Upper third | 80 (24.8) | 69 (27.2) | 11 (15.9) | *0.033* |
| Middle third | 108 (33.4) | 88 (34.6) | 20 (29.0) | |
| Lower third | 135 (41.8) | 97 (38.2) | 38 (55.1) | |
| Tumor size (cm) | | | | |
| Mean ± SD | 2.5±1.5 | 2.3±1.4 | 3.1±1.8 | *0.002* |
| Median | 2.0 | 2.0 | 2.6 | – |
| Main gross pattern | | | | |
| IIa | 24 (7.4) | 19 (7.5) | 5 (7.2) | 0.570 |
| IIb | 64 (19.8) | 54 (21.3) | 10 (14.5) | |
| IIc | 231 (71.5) | 178 (70.1) | 53 (76.8) | |
| III | 4 (1.2) | 3 (1.2) | 1 (1.4) | |
| UL | | | | |
| – | 195 (60.4) | 163 (64.2) | 32 (46.4) | *0.008* |
| + | 128 (39.6) | 91 (35.8) | 37 (53.6) | |

The italic numbers are P values with statistical significance (<0.05). –, absence; +, presence. BMI, body mass index; UL, ulcerative lesion.

**Table 2** Pathological characteristics (%)

| Data source | Total | Node metastasis | | P value |
| --- | --- | --- | --- | --- |
| | | Negative | Positive | |
| Biopsy | | | | |
| UDC grade* | | | | |
| U1 | 44 (13.6) | 35 (13.8) | 9 (13.0) | 0.518 |
| U2 | 43 (13.3) | 31 (12.2) | 12 (17.4) | |
| U3 | 236 (73.1) | 188 (74.0) | 48 (69.6) | |
| SRC grade** | | | | |
| S0 | 177 (54.8) | 136 (53.5) | 41 (59.4) | 0.415 |
| S1 | 146 (45.2) | 118 (46.5) | 28 (40.6) | |
| Surgery | | | | |
| UDC grade* | | | | |
| U1 | 24 (7.4) | 22 (8.7) | 2 (2.9) | 0.280 |
| U2 | 92 (28.5) | 72 (28.3) | 20 (29.0) | |
| U3 | 207 (64.1) | 160 (63.0) | 47 (68.1) | |
| SRC grade** | | | | |
| S0 | 169 (52.3) | 136 (53.5) | 33 (47.8) | 0.418 |
| S1 | 154 (47.7) | 118 (46.5) | 36 (52.2) | |
| Depth of invasion | | | | |
| LP | 113 (35.0) | 103 (40.6) | 10 (14.5) | <0.001 |
| MM | 52 (16.1) | 43 (16.9) | 9 (13.0) | |
| SM | 158 (48.9) | 108 (42.5) | 50 (72.5) | |
| LVI | | | | |
| – | 284 (87.9) | 238 (93.7) | 46 (66.7) | <0.001 |
| + | 39 (12.1) | 16 (6.3) | 23 (33.3) | |

*, r=0.523, P<0.001; **, r=0.491, P<0.001. The italic numbers are P values with statistical significance (<0.05). –, absence; +, presence. SRC, signet ring cell; LP, lamina propria; MM, muscularis mucosa; SM, submucosa; LVI, lymphovascular invasion.

## *Risk analysis*

### Preoperative parameters

In univariate analyses, including demographic, clinical and histological features of biopsy specimens: lower BMI, the presence of melena, larger tumor size and the lower third stomach were the risk factors of LNM (*Figure 2A*); lower BMI, the presence of melena, larger tumor size and the lower third stomach were the risk factors of LVI (*Figure 2B*); larger tumor size and the presence of UL were the risk factors of SM invasion (*Figure 2C*). While, the presence of SRC (S1) was the protective factor of LVI (*Figure 2B*);

the SRC grade of S1 and the UDC grade of U3 were the protective factors of SM invasion (*Figure 2C*); in T1a patients, the gender of female and the middle third stomach were the protective factors of muscularis mucosa (MM) invasion (*Figure 2D*). The detailed data of the univariate analyses could be found in *Tables S1-S4*.

In corresponding multivariate analyses: lower BMI, the presence of melena/UL, larger tumor size and the lower third stomach were the independent risk factors of LNM; lower BMI, the presence of melena and larger tumor size were the independent risk factors of LVI; larger tumor size
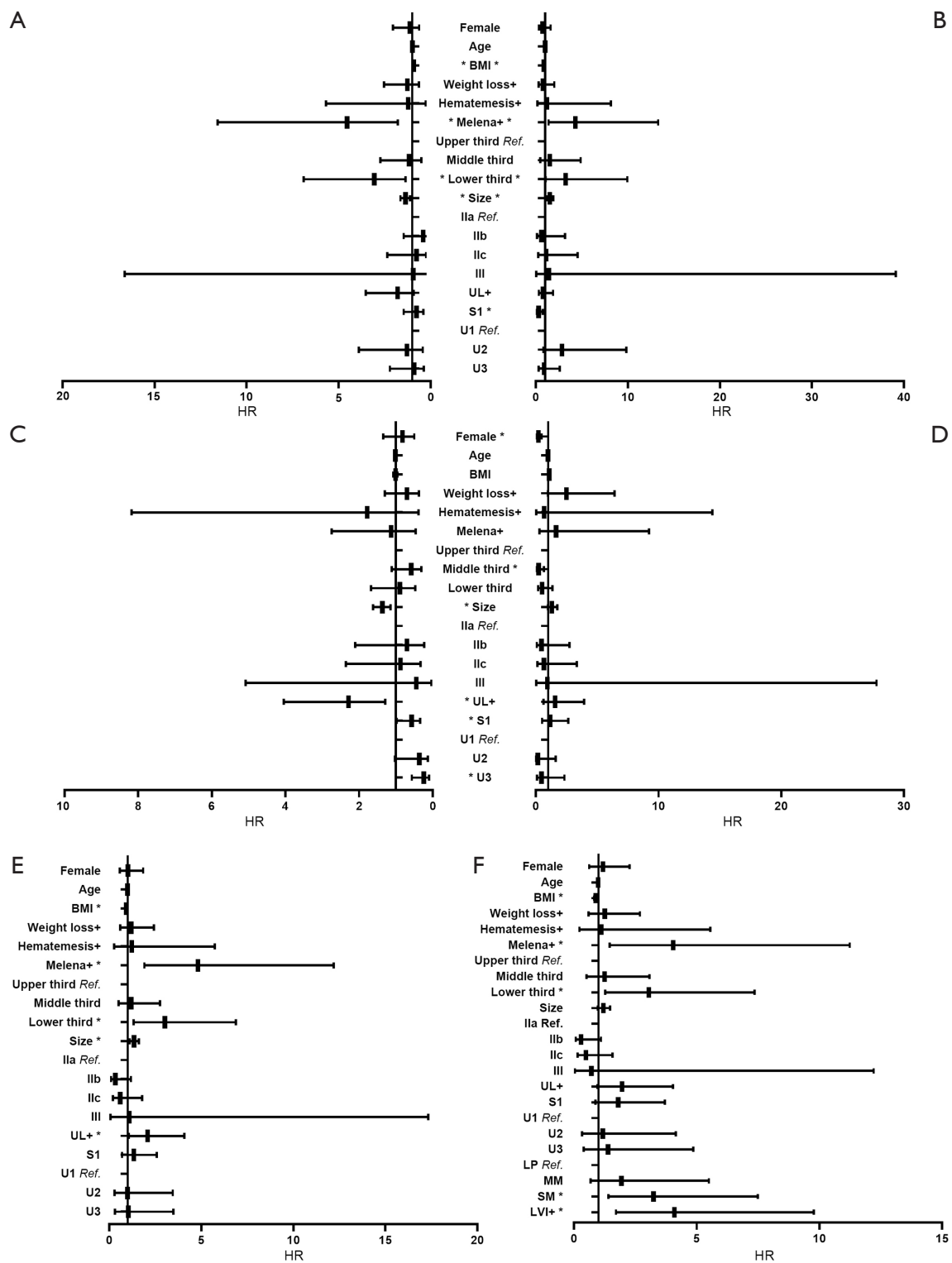
**Figure 2** Univariate analysis of pre- and postoperative data. **(A)** The univariate analysis for LNM; (B) the univariate analysis for LVI; (C) the univariate analysis for SM invasion; (D) the univariate analysis for MM invasion in T1a patients; (E) the univariate analysis for LNM; (F) the univariate analysis for LNM with data including LVI and depth of invasion. *, statistical significance. BMI, body mass index; UL, ulcerative lesion; LVI, lymphovascular invasion; Ref., reference; HR, hazard ratio.

**Table 3** Preoperative independent risk factors of tumor invasion, LVI and LNM

| Clinicopathological characteristics | Factors | Variables | P value | HR | 95% CI | |
|---|---|---|---|---|---|---|
| | | | | | Lower | Upper |
| LNM | BMI | CV | <0.001 | 0.878 | 0.846 | 0.912 |
| | Melena | + | <0.001 | 4.669 | 2.052 | 10.621 |
| | Tumor size | CV | 0.007 | 1.259 | 1.065 | 1.488 |
| | Tumor location | Upper third | Ref. | 1 | – | |
| | | Middle third | 0.759 | 1.137 | 0.502 | 0.759 |
| | | Lower third | 0.012 | 2.666 | 1.238 | 0.012 |
| | UL | + | 0.020 | 1.970 | 1.114 | 3.483 |
| LVI | BMI | CV | <0.001 | 0.882 | 0.853 | 0.911 |
| | Melena | + | 0.010 | 3.263 | 1.327 | 8.023 |
| | Tumor size | CV | <0.001 | 1.488 | 1.220 | 1.816 |
| | SRC | S1 | 0.001 | 0.247 | 0.108 | 0.565 |
| SM invasion | Tumor size | CV | <0.001 | 1.389 | 1.197 | 1.613 |
| | UL | + | <0.001 | 2.556 | 1.574 | 4.152 |
| | UDC grade | U1 | Ref. | 1 | – | |
| | | U2 | 0.033 | 0.424 | 0.193 | 0.033 |
| | | U3 | <0.001 | 0.291 | 0.163 | <0.001 |
| | SRC grade | S1 | 0.030 | 0.565 | 0.336 | 0.948 |
| MM invasion* | Sex | Female | <0.001 | 0.232 | 0.135 | 0.399 |

*, in T1a patients. The italic numbers are P values with statistical significance (<0.05). +, presence. CV, continuous variable; Ref., reference; BMI, body mass index; UL, ulcerative lesion; LVI, lymphovascular invasion; SM, submucosa; MM, muscularis mucosa; UDC, undifferentiated component; SRC, signet ring cell.

and the presence of UL were the independent risk factors of SM invasion. However, the SRC grade of S1 were the independent protective factors of LVI; the UDC grade of U2/U3 and the SRC grade of S1 were the independent protective factors of SM invasion; in T1a patients, the gender of female was the independent protective factor of MM invasion (*Table 3*).

## Postoperative parameters

In univariate analysis, including the same demographic, clinical but histological features of surgical specimens: lower BMI, the presence of melena/UL, larger tumor size and the lower third stomach were the risk factors of LNM (*Figure 2E*); in multivariate analysis, all the risk factors mentioned above showed independent significance to LNM (*Table 4*).

When depth of invasion and LVI were added into the

covariates, lower BMI, the presence of melena/LVI, the lower third stomach and SM invasion were risk factors of LNM in univariate analysis (*Figure 2F*), and lower BMI, the presence of melena/LVI, the lower third stomach, the SRC grade of S1 and SM invasion showed independent significance in multivariate analysis (*Table 4*). The detailed data of the univariate analyses could be found in *Tables S5,S6*.

## Comparison of the models with pre- or postoperative risk factors

The performance of the predictive models was evaluated by ROC curves. In those with the preoperative independent risk factors, the model for LNM (*Figure 3A*) showed the second best performance, exceeded by that for LVI (*Figure 3B*) and followed by those for SM (*Figure 3C*)

Page 8 of 13

Zou et al. Preoperative risk factors of early gastric cancer

**Table 4** Postoperative independent risk factors of LNM

| Condition | Factors | Variables | P value | HR | 95% CI | |
| --- | --- | --- | --- | --- | --- | --- |
| | | | | | Lower | Upper |
| Without LVI & T data | BMI | CV | *<0.001* | 0.878 | 0.846 | 0.912 |
| | Melena | + | *<0.001* | 4.669 | 2.052 | 10.621 |
| | Tumor location | Upper third | Ref. | 1 | – | |
| | | Middle third | 0.759 | 1.137 | 0.502 | 2.577 |
| | | Lower third | *0.012* | 2.666 | 1.238 | 5.741 |
| | Tumor size | CV | *0.007* | 1.259 | 1.065 | 1.488 |
| | UL | + | *0.020* | 1.970 | 1.114 | 3.483 |
| With LVI & T data | BMI | CV | *<0.001* | 0.862 | 0.824 | 0.902 |
| | Melena | + | *0.001* | 4.244 | 1.751 | 10.283 |
| | Tumor location | Upper third | Ref. | 1 | – | |
| | | Middle third | 0.643 | 1.224 | 0.521 | 2.877 |
| | | Lower third | *0.016* | 2.659 | 1.196 | 5.911 |
| | SRC | S1 | *0.038* | 1.912 | 1.038 | 3.522 |
| | Depth of invasion | LP | *0.011* | 1 | – | |
| | | MM | 0.186 | 1.945 | 0.725 | 5.216 |
| | | SM | *0.003* | 3.294 | 1.508 | 7.197 |
| | LVI | + | *0.001* | 3.832 | 1.694 | 8.667 |

The italic numbers are P values with statistical significance (<0.05). +, presence. CV, continuous variable; Ref., reference; BMI, body mass index; UL, ulcerative lesion; LP, lamina propria; MM, muscularis mucosa; SM, submucosa; SRC, signet ring cell; LVI, lymphovascular invasion.

and MM (*Figure 3D*) invasion. In models with the postoperative independent risk factors excluding depth of invasion and LVI, the model for LNM (*Figure 3E*) showed the same performance as that with preoperative data, and when depth of invasion (*Figure 3F*), LVI (*Figure 3G*), or both of them (*Figure 3H*) were added into the covariates, the performance of predictive models increased mildly.

## Discussion

EGC-UDC is a special entity of EGC, of which biological behaviors were proved to be more malignant than that of differentiated EGC (19). To the best of our knowledge, this is the first study to explore the risk factors of tumor invasion and LNM by combing clinical data and the histology of biopsy specimens, revealing that the preoperative information was significantly associated with tumor invasion

and/or LNM, the predictive models based on which showed comparable performance with those using postoperative data.

Due to the limited representativeness of biopsy specimens, the significance of their histological features to tumor behaviors attracted less attention in the past. Ryu (20) and Takao *et al.* (21) found a relatively high discrepancy rate between biopsy and surgical specimens, and believed that the diagnosis from biopsy specimens was easy to be affected by the accuracy of sampling. However, in Park's (22) and Park (23) work, biopsy specimens showed an overall accuracy of 87.2% in the pathological diagnosis of total EGC, and a specificity of 100% in detecting EGC-UDC. In the present study, using our grading system, biopsy and surgical specimens showed a significant histological correlation, which provided us a foundation for evaluating the significance of biopsy specimens, and the same result of Logistic regression or the same performance of predictive models also demonstrated the comparable risk strength of
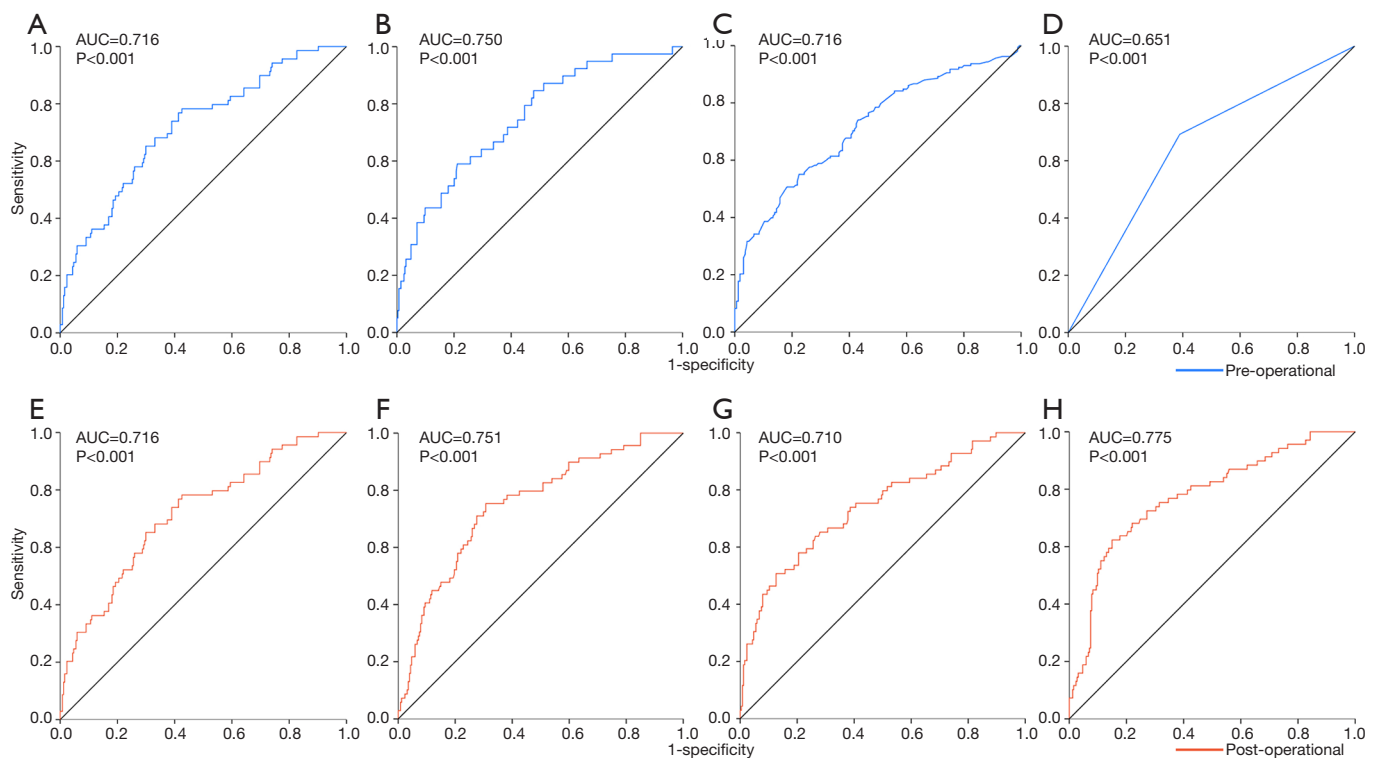
**Figure 3** The ROC curves of predictive models. (A) The ROC curve of the model for LNM with the preoperative independent risk factors; (B) the ROC curve of the model for LVI with the preoperative independent risk factors; (C) the ROC curve of the model for SM invasion with the preoperative independent risk factors; (D) the ROC curve of the model for MM invasion with the preoperative independent risk factors; (E) the ROC curve of the model for LNM with the postoperative independent risk factors screened out from the variates without depth of invasion and LVI; (F) the ROC curve of the model for LNM with the postoperative independent risk factors screened out from the variates containing depth of invasion; (G) the ROC curve of the model for LNM with the postoperative risk factors screened out from the variates containing LVI; (H) the ROC curve of the model for LNM with the postoperative risk factors screened out from the variates containing depth of invasion and LVI. AUC, areas under the ROC curve; LNM, lymph node metastasis; LVI, lymphovascular invasion; MM, muscularis mucosa.

biopsy and surgical specimens in histology.

In the present study, although two grading systems, UDC and SRC, were used to grab the histological features of biopsy specimens, no significant correlation was found between the histology and LNM. This was consistent with Kang's finding (24) that the different histological types of biopsy specimens were not correlated with LNM. But in Shim's work (25), moderately differentiated histology in biopsy specimens was reported to be relevant to LNM in "mixed histology predominantly of differentiated type EGC". Besides, the irrelevance between histology and LNM was also observed in surgical specimens here, which differed from some previous findings. Imamura (16) and Kim (26) revealed that the mixed-SRC histological type

were independent predictors of LNM in EGC-UDC, and the presence of SRC was an independent favorable prognostic factor. Horiuchi *et al.* (17) showed that the curative resection rate was significantly higher in pure SRCC than in mixed poorly differentiated GC. Jin *et al.* (27) also found that patients with SRCC had the lowest LNM rate than those with poorly differentiated or even moderately differentiated GC. Back to our study, the irrelevance between histology and LNM might be caused by adding clinical variables into the models, such as BMI and specific symptoms, which showed strong and persistent risk strength in the models for LNM, LVI, and/or SM invasion.

Nevertheless, it is worth noting that the presence of SRC

Page 10 of 13

Zou et al. Preoperative risk factors of early gastric cancer

in biopsy specimens significantly lower the risk of LVI in EGC-UDC independently. In numerous previous studies, as well as ours, LVI was closely correlated with node metastasis (19,28,29) or early distant metastasis (30), and functioned as an strong risk factor of LNM in EGC (odds ratio: 3.1–290.7) (10,16,24,26,31). However, rare studies had focused on the prediction of LVI with preoperative data in the past. As far as we have known, Kang *et al.* (24) were the first team to analyze the relationship between LVI and the histological types of biopsy specimens, and reported that, among poorly differentiated, SRCC (SRC >50%), poorly differentiated with SRCC (SRC ≤50%), mixed type, medullary type and poorly cohesive carcinoma (PCC), PCC and poorly differentiated/mixed differentiated histological type increased the risk of LVI independently. However, in our study, only the presence of SRC (grade S1) was identified as an independent protective factor of LVI. The difference between Kang and our results might be due to the different classifications of histological features: Kang *et al.* (24) defined a histological type of a tumor according to the tumor component exceeding 50%, while, in our study, UDC were graded in more detail and SRC was graded into two tiers (because of the relatively small sample size, the proportion of SRC was not subgrouped). The protective effect of SRC might be partially explained by the low microvessel density (MVD) in EGC with SRC, which was observed by imaging approach before (32). Intriguing, in the risk analysis using postoperative UDC and SRC grades, as well as the depth of invasion and LVI, the presence of SRC turned to a risk factor of LNM. Given the fact that more patients were found with SRC in surgical specimens than in biopsy specimens, SRC might function in a location-dependent manner, as only the adequate sampling of surgical specimens could reveal the SRC located in deep layers where SRC behaved differently from those in superficial layers (32). In addition, another explanation for the opposite roles of SRC in LVI and LNM might be the way SRC increased the risk of LNM, which might be not by increasing LVI.

Another dependent variable correlated with the histology of biopsy specimens was SM invasion, which is significant to the curability of ESD for tumors of expanded indication, including EGC-UDC (8). However, in practice nowadays, the depth of invasion is only concluded from endoscopic observation, CT images or lifting sign, causing SM invasion could be found in nearly 8% of ESD specimens (33). Our finding based on the histological features of biopsy specimens might provide a new direction for the evaluation of the depth of invasion, although, the performance of the predictive model was not excellent so far, and identifying more preoperative risk factors might be an effective way to improve the model performance in the future. Further, both UDC and SRC played protective roles in SM invasion here, combing the fact that the histological features of either biopsy or surgical specimens were not risk factors of LNM in our study, the traditional opinion on the equivalence between malignancy and undifferentiation was challenged. This situation pushes us to seek novel approaches to predict tumor behaviors with higher reliability. Integrated molecular classification of EGC might be a promising way, since some immunophenotypes (27), mi-RNA (34) and long noncoding RNAs (35), *et al.*, rather than the morphological differentiation alone, had been found relevant to LNM in EGC.

The role of BMI in EGC seems to be controversial so far. It was demonstrated that high BMI (≥25 kg/m²) increased the risk of EGC or mucosal dysplasia (36), with insulin resistance playing an important role (37). But other studies showed that lower BMI was associated with metachronous GC (38) and worse survival of EGC patients (39). Our analysis identified BMI as a protective factor of both LNM and LVI in EGC-UDC, which was consistent with previous findings that BMI was inversely associated with tumor size, tumor depth, LNM and tumor stage in early or advanced GC (40,41). To explore whether the controversy mentioned above is due to patient grouping, we evaluated the role of BMI in underweight (<18.5 kg/m²), normal (18.5–24 kg/m²) or overweight (≥24 kg/m²) patients separately, and found that the protective effect were persistent in all the three groups (*Table S7*). These results indicated that the different influences of BMI might be decided by different phases (before or after the occurrence of EGC) rather than the different amount of lipid reserve, and once a tumor starts, with excessive energy consumption, sufficient energy storage might be essential for anti-tumor response and a better outcome.

To explore the significance of clinical manifestation, weight loss, melena and hematemesis were selected as representative symptoms in this study. To our surprise, the presence of melena was firstly found increasing the risk of LNM and LVI by 4.898 and 4.606 times in EGC-UDC respectively. It was found that overt bleeding (hematemesis, melena or hematochezia) was an index symptom of malignancy (42), and occurred in 17.4% of EGC (43). Some scholars tried to screen EGC by detecting upper gastrointestinal tract bleeding (44,45), but the relatively

low specificity confines its application. However, in our study, the importance of melena opens a new gate for stomach bleeding detection, not only to screen EGC, but also to evaluate the risk of LNM and LVI. Further, as a risk factor of LVI and a parallel risk factor of LNM with LVI, melena might play a multifaced role in the progression of EGC-UDC. Firstly, it was relevant to the peri-tumoral vascular injury caused by tumor invasion in LVI; secondly, it might imply the nature of the tumors, since melena could be independent from LVI to increase the risk of LNM. In previous studies, the bleeding of GC was found closely associated with the vascular composition of tumors, according to which, GC could be categorized into hypo-vascular or hyper-vascular types (43), and the latter was proved to significantly correlates with LNM (46). So, given the parallel relationship between melena and LVI in the multivariate analysis for LNM, patients with melena might harbored hyper-vascular tumors, and the bleeding was not only due to the peri-tumoral vascular injury, but also related to the hypervascularity of these tumors. Thirdly, it is also possible that melena was an accompanied symptom which was not led by tumor invasion directly but induced through other unknown mechanisms, such as a weaker gastric mucosal barrier associated with some specific microbiota found in the past (47). However, more studies are needed to clarify the speculation above.

There were several limitations in our study. Firstly, due to the requirement of both surgical and matched preoperative biopsy data, although cases from three centers have been enrolled, the eligible cohort size was relatively small, which limited the statistical analysis, especially for EGC with different proportion of SRC. Secondly, the ability of the three symptoms selected to represent the clinical manifestation was relatively weak, and more attention should be paid to the importance of clinical data, since melena was identified as a strong independent risk factor of LNM and LVI. Thirdly, the role of gastric microbe was not involved, which is worth exploring in the future due to its role in tumorigenesis and progression of GC.

In summary, the preoperative data, including the histological features of biopsy specimens and clinical data were significantly associated with tumor invasion and/ or LNM, which deserves more attention. The predictive models with pre- or postoperative histological features showed comparable performance. Further studies based on a larger sample size will be needed to verify or even expand the significance of preoperative risk factors and build more efficient models to identify EGC-UDC patients eligible for

ESD treatment.

## Footnote

*Conflicts of Interest:* The authors have no conflicts of interest to declare.

*Ethical Statement:* The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was approved by the Institutional Review Board of Second Affiliated Hospital of Zhejiang University (2019-ERR-446), Union Hospital of Fujian Medical University (2020KY005) and Second Affiliated Hospital of Fujian Medical University (2019-SAHFMER-208). Patient consent was waived by the institutional review boards, as this study was retrospective and patients' information was protected by blind method.

## 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.

Page 12 of 13

Zou et al. Preoperative risk factors of early gastric cancer

2. Abdelfatah MM, Barakat M, Lee H, et al. The incidence of lymph node metastasis in early gastric cancer according to the expanded criteria in comparison with the absolute criteria of the Japanese Gastric Cancer Association: a systematic review of the literature and meta-analysis. Gastrointest Endosc 2018;87:338-47.

3. ZeLong Y, ZhenYu C, JianHai L, et al. Influence of Tumor Location on Lymph Node Metastasis and Survival for Early Gastric Cancer: a Population-Based Study. J Gastrointest Surg 2019;doi: 10.1007/s11605-019-04367-x.

4. Huang Q, Fang C, Shi J, et al. Differences in Clinicopathology of Early Gastric Carcinoma between Proximal and Distal Location in 438 Chinese Patients. Sci Rep 2015;5:13439.

5. Ono H, Yao K, Fujishiro M, et al. Guidelines for endoscopic submucosal dissection and endoscopic mucosal resection for early gastric cancer. Dig Endosc 2016;28:3-15.

6. Fang C, Shi J, Sun Q, et al. Risk factors of lymph node metastasis in early gastric carcinomas diagnosed by WHO criteria in 379 Chinese patients. J Dig Dis 2016;17:526-37.

7. Takizawa K, Ono H, Kakushima N, et al. Risk of lymph node metastases from intramucosal gastric cancer in relation to histological types: how to manage the mixed histological type for endoscopic submucosal dissection. Gastric Cancer 2013;16:531-6.

8. Japanese Gastric Cancer Association. Japanese gastric cancer treatment guidelines 2014 (ver. 4). Gastric Cancer 2017;20:1-19.

9. Wang ZK, Lin JX, Li P, et al. Higher Risk of Lymph Node Metastasis in Young Patients with Early Gastric Cancer. J Cancer 2019;10:4389-96.

10. Du MZ, Gan WJ, Yu J, et al. Risk factors of lymph node metastasis in 734 early gastric carcinoma radical resections in a Chinese population. J Dig Dis 2018;19:586-95.

11. Hu Q, Dekusaah R, Cao S, et al. Risk Factors of Lymph Node Metastasis in Patients with Early Pure and Mixed Signet Ring Cell Gastric Carcinomas. J Cancer 2019;10:1124-31.

12. Hirasawa T, Gotoda T, Miyata S, et al. Incidence of lymph node metastasis and the feasibility of endoscopic resection for undifferentiated-type early gastric cancer. Gastric Cancer 2009;12:148-52.

13. Nakamura R, Omori T, Mayanagi S, et al. Risk of lymph node metastasis in undifferentiated-type mucosal gastric carcinoma. World J Surg Oncol 2019;17:32.

14. Zhao X, Cai A, Xi H, et al. Predictive Factors for Lymph Node Metastasis in Undifferentiated Early Gastric Cancer:
 a Systematic Review and Meta-analysis. J Gastrointest Surg 2017;21:700-11.

15. Bang CS, Yang YJ, Lee JJ, et al. Endoscopic Submucosal Dissection of Early Gastric Cancer with Mixed-Type Histology: A Systematic Review. Dig Dis Sci 2020;65:276-91.

16. Imamura T, Komatsu S, Ichikawa D, et al. Early signet ring cell carcinoma of the stomach is related to favorable prognosis and low incidence of lymph node metastasis. J Surg Oncol 2016;114:607-12.

17. Horiuchi Y, Fujisaki J, Yamamoto N, et al. Mixed poorly differentiated adenocarcinoma in undifferentiated-type early gastric cancer predicts endoscopic noncurative resection. Gastric Cancer 2018;21:689-95.

18. The Paris endoscopic classification of superficial neoplastic lesions: esophagus, stomach, and colon: November 30 to December 1, 2002. Gastrointest Endosc 2003;58:S3-43.

19. Gu L, Chen M, Khadaroo PA, et al. A Risk-Scoring Model for Predicting Lymph Node Metastasis in Early Gastric Cancer Patients: a Retrospective Study and External Validation. J Gastrointest Surg 2018;22:1508-15.

20. Ryu DG, Choi CW, Kang DH, et al. Predictive factors to diagnosis undifferentiated early gastric cancer after endoscopic submucosal dissection. Medicine (Baltimore) 2017;96:e8044.

21. Takao M, Kakushima N, Takizawa K, et al. Discrepancies in histologic diagnoses of early gastric cancer between biopsy and endoscopic mucosal resection specimens. Gastric Cancer 2012;15:91-6.

22. Park JY, Ryu KW, Eom BW, et al. Proposal of the surgical options for primary tumor control during sentinel node navigation surgery based on the discrepancy between preoperative and postoperative early gastric cancer diagnoses. Ann Surg Oncol 2014;21:1123-9.

23. Park HK, Lee KY, Yoo MW, et al. Mixed Carcinoma as an Independent Prognostic Factor in Submucosal Invasive Gastric Carcinoma. J Korean Med Sci 2016;31:866-72.

24. Kang SH, Kim JS, Moon HS, et al. Signet ring cell carcinoma of early gastric cancer, is endoscopic treatment really risky? Medicine (Baltimore) 2017;96:e7532.

25. Shim CN, Chung H, Park JC, et al. Early gastric cancer with mixed histology predominantly of differentiated type is a distinct subtype with different therapeutic outcomes of endoscopic resection. Surg Endosc 2015;29:1787-94.

26. Kim YH, Park JH, Park CK, et al. Histologic purity of signet ring cell carcinoma is a favorable risk factor for lymph node metastasis in poorly cohesive, submucosainvasive early gastric carcinoma. Gastric Cancer

2017;20:583-90.

27. Jin EH, Lee DH, Jung SA, et al. Clinicopathologic factors and molecular markers related to lymph node metastasis in early gastric cancer. World J Gastroenterol 2015;21:571-7.

28. Hwang CS, Ahn S, Lee BE, et al. Risk of lymph node metastasis in mixed-type early gastric cancer determined by the extent of the poorly differentiated component. World J Gastroenterol 2016;22:4020-6.

29. Barreto SG, Windsor JA. Redefining early gastric cancer. Surg Endosc 2016;30:24-37.

30. Bao J, Qu G, Fu W, et al. Clinicopathological features of gastric adenocarcinoma patients with metachronous distant metastasis. Tumour Biol 2015;36:6375-82.

31. Zhao B, Zhang J, Zhang J, et al. Risk Factors Associated with Lymph Node Metastasis for Early Gastric Cancer Patients Who Underwent Non-curative Endoscopic Resection: a Systematic Review and Meta-analysis. J Gastrointest Surg 2019;23:1318-28.

32. Chen J, Cai R, Ren G, et al. Differences in clinicopathological characteristics and computed tomography findings between signet ring cell carcinoma. Cancer Med 2018;7:1160-9.

33. Kim JS, Kang SH, Moon HS, et al. Clinical outcome after endoscopic submucosal dissection for early gastric cancer of absolute and expanded indication. Medicine (Baltimore) 2017;96:e6710.

34. Ma M, Lu S, Liu Y, et al. Identification and external validation of a novel miRNA signature for lymph node metastasis prediction in submucosal-invasive gastric cancer patients. Cancer Med 2019;Epub ahead of print.

35. Lu Q, Yu T, Ou X, et al. Potential lncRNA diagnostic biomarkers for early gastric cancer. Mol Med Rep 2017;16:9545-52.

36. Kim HJ, Kim N, Kim HY, et al. Relationship between body mass index and the risk of early gastric cancer and dysplasia regardless of Helicobacter pylori infection. Gastric Cancer 2015;18:762-73.

37. Kwon HJ, Park MI, Park SJ, et al. Insulin Resistance

Is Associated with Early Gastric Cancer: A Prospective Multicenter Case Control Study. Gut Liver 2019;13:154-60.

38. Kim JL, Kim SG, Kim J, et al. Clinical Outcomes of Metachronous Gastric Cancer after Endoscopic Resection for Early Gastric Cancer. Gut Liver 2020;14:190-8.

39. Matsunaga T, Saito H, Osaki T, et al. Use of Body Mass Index to Predict the Prognosis of Patients with Remnant Gastric Cancer. Yonago Acta Med 2017;60:126-32.

40. Feng F, Zheng GZ, Guo XH, et al. Impact of body mass index on surgical outcomes of gastric cancer. BMC Cancer 2018;18:151.

41. Komatsu S, Kosuga T, Kubota T, et al. Preoperative Low Weight Affects Long-term Outcomes Following Curative Gastrectomy for Gastric Cancer. Anticancer Res 2018;38:5331-7.

42. Schatz RA, Rockey DC. Gastrointestinal Bleeding Due to Gastrointestinal Tract Malignancy: Natural History, Management, and Outcomes. Dig Dis Sci 2017;62:491-501.

43. Wang L, Wang XA, Hao JQ, et al. Long-term outcomes after radical gastrectomy in gastric cancer patients with overt bleeding. World J Gastroenterol 2015;21:13316-24.

44. Liu H, Qiao P, Wu X, et al. A smart capsule system of gastric occult blood detection. Biomed Mater Eng 2014;24:519-28.

45. Harewood GC, McConnell JP, Harrington JJ, et al. Detection of occult upper gastrointestinal tract bleeding: performance differences in fecal occult blood tests. Mayo Clin Proc 2002;77:23-8.

46. Hong WG, Ko YS, Pyo JS. Clinicopathological significance and prognostic role of microvessel density in gastric cancer: A meta-analysis. Pathol Res Pract 2017;213:1459-63.

47. Liu X, Shao L, Liu X, et al. Alterations of gastric mucosal microbiota across different stomach microhabitats in a cohort of 276 patients with gastric cancer. EBioMedicine 2019;40:336-48.

## Supplementary

**Table S1** Multivariate analysis for LNM with preoperative data

| Factors | Univariate analysis | | | |
|---|---|---|---|---|
| | P | HR | 95% CI | |
| | | | Lower | Upper |
| Sex | | | | |
| Male | Ref. | 1 | – | |
| Female | 0.684 | 1.131 | 0.626 | 2.041 |
| Age | | | | |
| CV | 0.177 | 0.985 | 0.963 | 1.007 |
| BMI | | | | |
| CV | 0.014 | 0.919 | 0.859 | 0.983 |
| Weight loss | | | | |
| – | Ref. | 1 | – | |
| + | 0.511 | 1.264 | 0.629 | 2.538 |
| Hematemesis | | | | |
| – | Ref. | 1 | – | |
| + | 0.806 | 1.214 | 0.259 | 5.694 |
| Melena | | | | |
| – | Ref. | 1 | – | |
| + | 0.002 | 4.535 | 1.778 | 11.570 |
| Location | | | | |
| Upper third | Ref. | 1 | – | |
| Middle third | 0.717 | 1.170 | 0.500 | 2.735 |
| Lower third | 0.007 | 3.063 | 1.360 | 6.897 |
| Size | | | | |
| CV | 0.002 | 1.357 | 1.123 | 1.639 |
| Gross pattern | | | | |
| IIa | Ref. | 1 | – | |
| IIb | 0.166 | 0.398 | 0.109 | 1.464 |
| IIc | 0.653 | 0.774 | 0.254 | 2.363 |
| III | 0.972 | 0.950 | 0.054 | 16.623 |
| UL | | | | |
| – | Ref. | 1 | – | |
| + | 0.088 | 1.797 | 0.916 | 3.526 |
| UDC grade | | | | |
| U1 | Ref. | 1 | – | |
| U2 | 0.655 | 1.288 | 0.425 | 3.909 |
| U3 | 0.823 | 0.903 | 0.369 | 2.209 |
| SRC grade | | | | |
| S0 | Ref. | 1 | – | |
| S1 | 0.419 | 0.765 | 0.400 | 1.464 |

–, absence; +, presence. CV, continuous variable; Ref., reference; BMI, body mass index; UL, ulcerative lesion; UDC, undifferentiated component; SRC, signet ring cell.

**Table S2** Multivariate analysis for LVI with preoperative data

| Factors | Univariate analysis | | | |
| --- | --- | --- | --- | --- |
| | P | HR | 95% CI | |
| | | | Lower | Upper |
| Sex | | | | |
| Male | Ref. | 1 | – | |
| Female | 0.426 | 0.725 | 0.328 | 1.601 |
| Age | | | | |
| CV | 0.791 | 0.996 | 0.968 | 1.025 |
| BMI | | | | |
| CV | 0.001 | 0.865 | 0.792 | 0.944 |
| Weight loss | | | – | |
| – | Ref. | 1 | – | |
| + | 0.598 | 0.774 | 0.300 | 2.001 |
| Hematemesis | | | | |
| – | Ref. | 1 | – | |
| + | 0.864 | 1.183 | 0.172 | 8.155 |
| Melena | | | | |
| – | Ref. | 1 | – | |
| + | 0.012 | 4.284 | 1.381 | 13.294 |
| Location | | | | |
| Upper third | Ref. | 1 | – | |
| Middle third | 0.481 | 1.519 | 0.474 | 4.867 |
| Lower third | 0.042 | 3.224 | 1.045 | 9.948 |
| Size | | | | |
| CV | 0.001 | 1.503 | 1.190 | 1.898 |
| Gross pattern | | | | |
| IIa | Ref. | 1 | – | |
| IIb | 0.558 | 0.612 | 0.119 | 3.161 |
| IIc | 0.894 | 1.101 | 0.267 | 4.535 |
| III | 0.843 | 1.401 | 0.050 | 39.111 |
| UL | | | | |
| – | Ref. | 1 | – | |
| + | 0.569 | 0.774 | 0.319 | 1.873 |
| UDC grade | | | | |
| U1 | Ref. | 1 | – | |
| U2 | 0.099 | 2.845 | 0.823 | 9.834 |
| U3 | 0.834 | 0.892 | 0.305 | 2.607 |
| SRC grade | | | | |
| S0 | Ref. | 1 | – | |
| S1 | 0.009 | 0.290 | 0.115 | 0.732 |

–, absence; +, presence. CV, continuous variable; Ref., reference; BMI, body mass index; UL, ulcerative lesion; UDC, undifferentiated component; SRC, signet ring cell.

**Table S3** Multivariate analysis for SM with preoperative data

| Factors | Univariate analysis | | | |
| --- | --- | --- | --- | --- |
| | P | HR | 95% CI | |
| | | | Lower | Upper |
| Sex | | | | |
| Male | Ref. | 1 | – | |
| Female | 0.429 | 0.819 | 0.499 | 1.343 |
| Age | | | | |
| CV | 0.221 | 1.012 | 0.993 | 1.031 |
| BMI | | | | |
| CV | 0.804 | 1.007 | 0.951 | 1.067 |
| Weight loss | | | – | |
| – | Ref. | 1 | – | |
| + | 0.260 | 0.701 | 0.377 | 1.301 |
| Hematemesis | | | | |
| – | Ref. | 1 | – | |
| + | 0.458 | 1.781 | 0.388 | 8.179 |
| Melena | | | | |
| – | Ref. | 1 | – | |
| + | 0.791 | 1.128 | 0.463 | 2.748 |
| Location | | | | |
| Upper third | Ref. | 1 | – | |
| Middle third | 0.101 | 0.585 | 0.309 | 1.110 |
| Lower third | 0.719 | 0.891 | 0.475 | 1.671 |
| Size | | | | |
| CV | 0.000 | 1.364 | 1.146 | 1.623 |
| Gross pattern | | | | |
| IIa | Ref. | 1 | – | |
| IIb | 0.523 | 0.698 | 0.231 | 2.104 |
| IIc | 0.796 | 0.877 | 0.326 | 2.359 |
| III | 0.512 | 0.442 | 0.038 | 5.083 |
| UL | | | | |
| – | Ref. | 1 | – | |
| + | 0.005 | 2.287 | 1.291 | 4.050 |
| UDC grade | | | | |
| U1 | Ref. | 1 | – | |
| U2 | 0.055 | 0.361 | 0.128 | 1.021 |
| U3 | 0.001 | 0.241 | 0.102 | 0.567 |
| SRC grade | | | | |
| S0 | Ref. | 1 | – | |
| S1 | 0.041 | 0.574 | 0.337 | 0.978 |

–, absence; +, presence. CV, continuous variable; Ref., reference; BMI, body mass index; UL, ulcerative lesion; UDC, undifferentiated component; SRC, signet ring cell.

**Table S4** Multivariate analysis for MM with preoperative data of T1a patients

| Factors | Univariate analysis | | | |
|---|---|---|---|---|
| | P | HR | 95% CI | |
| | | | Lower | Upper |
| Sex | | | | |
| Male | Ref. | 1 | – | |
| Female | 0.000 | 0.204 | 0.089 | 0.469 |
| Age | | | | |
| CV | 0.325 | 0.985 | 0.956 | 1.015 |
| BMI | | | | |
| CV | 0.099 | 1.081 | 0.985 | 1.187 |
| Weight loss | | | | |
| – | Ref. | 1 | – | |
| + | 0.057 | 2.497 | 0.972 | 6.417 |
| Hematemesis | | | | |
| – | Ref. | 1 | – | |
| + | 0.800 | 0.673 | 0.031 | 14.390 |
| Melena | | | | |
| – | Ref. | 1 | - | |
| + | 0.572 | 1.643 | 0.293 | 9.212 |
| Location | | | | |
| Upper third | Ref. | 1 | – | |
| Middle third | 0.006 | 0.223 | 0.077 | 0.646 |
| Lower third | 0.178 | 0.507 | 0.189 | 1.362 |
| Size | | | | |
| CV | 0.081 | 1.298 | 0.968 | 1.740 |
| Gross pattern | | | | |
| IIa | Ref. | 1 | – | |
| IIb | 0.392 | 0.457 | 0.076 | 2.741 |
| IIc | 0.620 | 0.664 | 0.132 | 3.339 |
| III | 0.970 | 0.938 | 0.032 | 27.767 |
| UL | | | | |
| – | Ref. | 1 | – | |
| + | 0.336 | 1.570 | 0.627 | 3.936 |
| UDC grade | | | | |
| U1 | Ref. | 1 | – | |
| U2 | 0.121 | 0.166 | 0.017 | 1.602 |
| U3 | 0.345 | 0.457 | 0.090 | 2.325 |
| SRC grade | | | | |
| S0 | Ref. | 1 | – | |
| S1 | 0.747 | 1.147 | 0.498 | 2.641 |

–, absence; +, presence. CV, continuous variable; Ref., reference; BMI, body mass index; UL, ulcerative lesion; UDC, undifferentiated component; SRC, signet ring cell.

**Table S5** Multivariate analysis for LNM with postoperative data (depth of invasion and LVI excluded)

| Factors | Univariate analysis | | | |
| --- | --- | --- | --- | --- |
| | P | HR | 95% CI | |
| | | | Lower | Upper |
| Sex | | | | |
| Male | Ref. | 1 | – | |
| Female | 0.960 | 1.015 | 0.561 | 1.839 |
| Age | | | | |
| CV | 0.301 | 0.988 | 0.966 | 1.011 |
| BMI | | | | |
| CV | 0.008 | 0.909 | 0.848 | 0.976 |
| Weight loss | | | | |
| – | Ref. | 1 | – | |
| + | 0.626 | 1.192 | 0.588 | 2.415 |
| Hematemesis | | | | |
| – | Ref. | 1 | – | |
| + | 0.807 | 1.214 | 0.257 | 5.734 |
| Melena | | | | |
| – | Ref. | 1 | – | |
| + | 0.001 | 4.815 | 1.900 | 12.199 |
| Location | | | | |
| Upper third | Ref. | 1 | – | |
| Middle third | 0.702 | 1.180 | 0.506 | 2.752 |
| Lower third | 0.009 | 3.020 | 1.324 | 6.887 |
| Size | | | | |
| CV | 0.003 | 1.334 | 1.104 | 1.612 |
| Gross pattern | | | | |
| IIa | Ref. | 1 | – | |
| IIb | 0.086 | 0.328 | 0.092 | 1.173 |
| IIc | 0.348 | 0.588 | 0.194 | 1.782 |
| III | 0.960 | 1.073 | 0.066 | 17.328 |
| UL | | | | |
| – | Ref. | 1 | – | |
| + | 0.034 | 2.076 | 1.057 | 4.077 |
| UDC grade | | | | |
| U1 | Ref. | 1 | – | |
| U2 | 0.994 | .995 | 0.287 | 3.447 |
| U3 | 0.965 | 1.028 | 0.303 | 3.484 |
| SRC grade | | | | |
| S0 | Ref. | 1 | – | |
| S1 | 0.396 | 1.331 | 0.688 | 2.573 |

–, absence; +, presence. CV, continuous variable; Ref., reference; BMI, body mass index, UL, ulcerative lesion; UDC, undifferentiated component; SRC, signet ring cell.

**Table S6** Multivariate analysis for LNM with postoperative data (containing depth of invasion and LVI)

| Factors | Univariate analysis | | | |
|---|---|---|---|---|
| | P | HR | 95% CI | |
| | | | Lower | Upper |
| Sex | | | | |
| Male | Ref. | 1 | – | |
| Female | 0.613 | 1.182 | 0.618 | 2.260 |
| Age | | | | |
| CV | 0.106 | 0.980 | 0.956 | 1.004 |
| BMI | | | | |
| CV | 0.003 | 0.889 | 0.822 | 0.962 |
| Weight loss | | | – | |
| – | Ref. | 1 | – | |
| + | 0.556 | 1.256 | 0.588 | 2.684 |
| Hematemesis | | | | |
| – | Ref. | 1 | – | |
| + | 0.903 | 1.106 | 0.220 | 5.560 |
| Melena | | | | |
| – | Ref. | 1 | – | |
| + | 0.007 | 4.043 | 1.455 | 11.237 |
| Location | | | | |
| Upper third | Ref. | 1 | – | |
| Middle third | 0.631 | 1.247 | 0.506 | 3.075 |
| Lower third | 0.013 | 3.053 | 1.266 | 7.360 |
| Size | | | | |
| CV | 0.094 | 1.192 | 0.970 | 1.464 |
| Gross pattern | | | | |
| IIa | Ref. | 1 | – | |
| IIb | 0.068 | 0.281 | 0.072 | 1.100 |
| IIc | 0.225 | 0.484 | 0.150 | 1.562 |
| III | 0.813 | 0.710 | 0.041 | 12.215 |
| UL | | | | |
| – | Ref. | 1 | – | |
| + | 0.069 | 1.957 | 0.949 | 4.034 |
| UDC grade | | | | |
| U1 | Ref. | 1 | – | |
| U2 | 0.813 | 1.165 | 0.327 | 4.149 |
| U3 | 0.608 | 1.388 | 0.396 | 4.857 |
| SRC grade | | | | |
| S0 | Ref. | 1 | – | |
| S1 | 0.110 | 1.801 | 0.876 | 3.702 |
| LVI | | | | |
| – | Ref. | 1 | – | |
| + | 0.002 | 4.087 | 1.709 | 9.769 |
| Depth of invasion | | | | |
| Lamina propria | Ref. | 1 | – | |
| Muscularis mucosa | 0.217 | 1.932 | 0.679 | 5.495 |
| Submucosa | 0.006 | 3.240 | 1.401 | 7.493 |

–, absence; +, presence. CV, continuous variable; Ref., reference; BMI, body mass index; UL, ulcerative lesion; UDC, undifferentiated component; SRC, signet ring cell; LVI, lymphovascular invasion.

**Table S7** The significant risk factors of LNM in patients with different BMI

| BMI (kg/m²) | Factors | Variables | P value | HR | 95% CI Lower | 95% CI Upper |
|---|---|---|---|---|---|---|
| <18.5 | BMI | CV | 0.026 | 0.938 | 0.886 | 0.992 |
| 18.5–24 | BMI | CV | <0.001 | 0.832 | 0.760 | 0.911 |
| | Melena | + | <0.001 | 8.635 | 3.217 | 23.173 |
| | Location | Upper third | Ref. | 1 | – | |
| | | Middle third | 0.531 | 1.396 | 0.491 | 3.968 |
| | | Lower third | 0.004 | 4.591 | 1.625 | 12.971 |
| | Gross pattern | IIa | Ref. | 1 | – | |
| | | IIb | 0.538 | 1.807 | 0.275 | 11.872 |
| | | IIc | 0.068 | 4.926 | 0.888 | 27.339 |
| | | III | 0.042 | 24.918 | 1.124 | 552.444 |
| ≥24 | BMI | CV | <0.001 | 0.866 | 0.812 | 0.924 |
| | Size | CV | 0.004 | 1.936 | 1.228 | 3.053 |
| | UL | + | 0.003 | 6.935 | 1.973 | 24.376 |

+, presence. CV, continuous variable; Ref., reference; BMI, body mass index; UL, ulcerative lesion.