Adjuvant chemoradiotherapy versus adjuvant chemotherapy for patients with N3 gastric cancer after D2/R0 resection: a retrospective study based on propensity score analyses

Meng-Long Zhou1,2,*
Wang Yang1,2,*
Ya-Qi Wang2,*
Miao Mo2,3
Ran Hu1,2
Yan Wang1,2
Jia-Ning Yang1,2
Gui-Chao Li1,2
Ya-Nong Wang2,4
Zhen Zhang1,2

1Department of Radiation Oncology, Fudan University Shanghai Cancer Center, Shanghai 200032, People’s Republic of China; 2Department of Oncology, Shanghai Medical College, Fudan University, Shanghai 200032, People’s Republic of China; 3Department of Cancer Prevention, Fudan University Shanghai Cancer Center, Shanghai, 200032, People’s Republic of China; 4Department of Gastric Surgery, Fudan University Shanghai Cancer Center, Shanghai 200032, People’s Republic of China

*These authors contributed equally to this work

Purpose: N3 gastric cancer (GC) is characterized by a heavy burden of lymph node metastasis and a high postoperative recurrence rate. The role of radiotherapy in this group of patients remains undetermined. The purpose of this study was to compare the effectiveness of adjuvant chemoradiotherapy (CRT) and adjuvant chemotherapy (ChT) for N3 GC after D2/R0 resection.

Patients and methods: From January 2004 to December 2015, patients with N3 GC in the database of Fudan University Shanghai Cancer Center were retrospectively reviewed. The eligible patients were enrolled in an adjuvant CRT group and an adjuvant ChT group. Four different methods based on a propensity score model were used to balance the baseline characteristics. Then, survival analyses between the two groups were performed in addition to patterns of recurrence and subgroup analyses.

Results: In total, 175 and 365 eligible patients were enrolled into the CRT and ChT groups, respectively. After balancing, the disease-free survival (DFS) of patients in the CRT group was significantly better than that of patients in the ChT group (p=0.021). Subgroup analyses showed that patients with N3a GC benefited from adjuvant CRT.

Conclusion: Compared with adjuvant ChT, adjuvant CRT can further improve the DFS of patients with N3 GC after D2/R0 resection. Patients with lymph node metastases should be further stratified when selecting patients for adjuvant CRT.

Keywords: stomach neoplasms, gastrectomy, chemoradiation, propensity score, survival analysis

Introduction

The incidence of gastric cancer (GC) has been generally declining for decades, but it is still the sixth most common cancer and the fourth most frequent cause of cancer-related death worldwide.1 The recently updated epidemiologic data in China show that 70.8% of cases of GC are locally advanced GC (LAGC). Therefore, how to appropriately treat these patients and improve their survival is a serious challenge for Chinese doctors and researchers.

There are now three available treatment strategies for LAGC including D2 surgery plus adjuvant chemotherapy (ChT), adjuvant chemoradiation (CRT) and perioperative ChT. However, patients with N3 GC, particularly N3b GC, which is characterized by a heavy burden of lymph node (LN) metastasis and...
high postoperative recurrence rate, are not often included in clinical studies. Thus, the optimal treatment strategy for these patients remains undetermined.

In addition, the role of adjuvant CRT for LAGC after D2/R0 surgery remains to be clarified. The subgroup analysis in the ARTIST trial indicated that GC patients with LN metastasis (pN+)) could benefit from adjuvant CRT after D2/R0 resection. However, patients with N3 GC accounted for no more than 30% of the total population in this study. The positive results of this analysis can likely be ascribed to the patients with N1-2 GC included in this subgroup.

There are several relevant retrospective studies that have yielded inconsistent results. A study conducted by Kilic et al, focused on patients with more than 16 LN metastases and found that compared with adjuvant ChT, adjuvant CRT did not yield further survival benefits. The subgroup analysis in the study at our center suggested the ineffectiveness of adjuvant CRT for patients with N3 GC. However, the subgroup analysis in another study conducted by Peng et al. indicated that patients with stage IIIC disease (including stage T4aN3, T4bN2 and T4bN3 disease) in the CRT group had significantly longer disease-free survival (DFS) and overall survival (OS) than those in the ChT group.

Thus, at present, there is a lack of randomized controlled trials (RCTs) that have evaluated if the addition of radiotherapy to adjuvant ChT can further benefit patients with N3 GC. In addition, the findings of several retrospective studies are inconsistent. The purpose of this study was to compare the effectiveness of adjuvant CRT and adjuvant ChT for N3 GC after D2/R0 resection to provide more evidence for treatment strategies for patients with N3 GC.

**Materials and methods**

**Patients**

All the consecutive patients who met the following inclusion criteria were first identified from the database: (1) D2/R0 gastrectomy performed between January 1st, 2004 and December 31st, 2015; (2) histologically confirmed adenocarcinoma of the stomach or gastroesophageal junction (GEJ); and (3) pathological confirmation of ≥7 metastatic LNs (pTanyN3 according to the 7th edition of the AJCC staging system).

The following patients were excluded: (1) patients who received single-agent ChT; (2) patients who received neoadjuvant ChT or CRT for a primary tumor; (3) patients who did not have detailed records of ChT; (4) patients who died within 30 days of surgery; (5) patients who were lost to follow-up immediately after discharge from the hospital; (6) patients who had peritoneal or distant metastases (stage IV disease) at the time of surgery; (7) patients who had R1 resection, which was defined as positive peritoneal cytology (CY1) or microscopically positive resection margins; (8) patients who had R2 resection, which was defined as gross residual tumors left behind after surgery; (9) patients with other pathological types of GC, such as squamous cell carcinoma or neuroendocrine tumors (NETs); (10) patients who underwent surgery due to gastric stump carcinoma; (11) patients who had double primary malignancies; (12) patients who were missing important clinicopathological data due to incomplete medical records; and (13) patients in the CRT group who did not finish the planned radiotherapy (RT).

**Treatment delivery**

Radiotherapy was given with 6 MV photons using either 3-dimensional CRT or intensity-modulated radiation therapy (IMRT). Patients were treated with 25–28 fractions of 1.8 Gy for a total dose of 45–50.4 Gy (5 fractions/week). The clinical target volume (CTV) encompassed the preoperative tumor extension, tumor bed, anastomosis site, and regional draining LNs. The planning target volume (PTV) margin was 0.5–1.0 cm considering the individual uncertainties. The remnant stomach has not been routinely included within the radiation field since 2008 due to the results of a study from Korea.

Concurrent ChT regimens included (1) a continuous intravenous infusion of 225 mg/m² of 5-fluorouracil (5-FU) for 120 hrs each week, and (2) 625 mg/m² of capecitabine twice daily from day 1–5 weekly or tegafur gimeracil oteracil potassium capsules (S1) 40 mg/m² twice daily from day 1–5 weekly.

The patients in the CRT group received one or two courses of adjuvant ChT followed by CRT and four to five additional subsequent courses of ChT. The patients in the ChT group received six to eight courses of ChT. The ChT regimens mostly consisted of combination therapies involving 5-fluorouracil (5-FU) or oral fluorouracil derivatives as the backbone with the addition of oxaliplatin, epirubicin or taxanes.

**Follow-up**

After completion of primary treatment, patients were followed every 3 months for the first 2 years, then every 6 months until the fifth year, and yearly thereafter. Follow-
up examinations included a complete history and physical examination, measurement of serum tumor markers, CT scans of the chest, abdomen and pelvis each time and endoscopy each year. Disease recurrence was documented by biopsy or cytology reports or by clinical evidence on radiographic studies. Toxicities were assessed using the Common Terminology Criteria for Adverse Events (CTCAE) 4.0.

**Definition of recurrence**
Recurrences were categorized as local recurrence (LR), regional recurrence (RR), peritoneal metastasis (PM) or distant metastasis (DM). LR was defined as any relapse at the site of anastomosis, remnant stomach, duodenal stump or tumor bed. RR was defined as recurrence involving the regional LNs within the RT field or hypothetical RT field of patients in the ChT group. Recurrence inside the peritoneal cavity was considered PM and included carcinomatosis of the colorectum, ovaries, peritoneum, and ureters. DM was considered metastases involving the intra-abdominal organs (liver, adrenal gland, etc.) or extra-abdominal organs (lung, bone, brain, etc.) and LNs (Virchow’s node, etc.).

**Predictor variables**
Clinicopathological variables considered in the analyses were classified into patient, clinical, and histopathological variables. Patient variables included sex and age. Clinical variables included the primary location of the tumor and the regimens and courses of adjuvant ChT. Histopathological variables included the World Health Organization (WHO) histology and grade of the tumor, number of positive LNs and dissected LNs, invasion depth, nodal status and AJCC stage of the tumor.

The courses of ChT were transformed into a dichotomous variable (complete vs incomplete) based on the number of courses. According to the design of the ARTIST trial, patients in the CRT group who received four or more courses of ChT or patients in the ChT group who received six or more courses of ChT were regarded as receiving complete adjuvant ChT. If they received fewer courses of treatment, they were considered to have received incomplete adjuvant ChT.

The adjuvant CRT and ChT groups were compared with respect to the patient, clinical and histopathological variables outlined. Continuous variables are reported as the mean (± SD) or median (with range), as appropriate, and categorical variables are reported as numbers and percentages. Variables were compared using the χ² test or independent sample t-test/Wilcoxon rank sum test, as appropriate. Univariable logistic regression models were used to estimate crude ORs and 95% CIs to evaluate the association between each variable and treatment status (adjuvant CRT or ChT).

**Statistical analysis**
The primary objective of this study was to compare the DFS between the two groups. The secondary objective was to compare the OS and determine the effectiveness of adjuvant CRT in clinically relevant patient subsets.

DFS was calculated from surgery until the first locoregional, distant or peritoneal recurrence (event) or the day of last follow-up without recurrence (censored). OS was calculated from surgery until death of any cause (event) or last follow-up contact (censored). Locoregional recurrence-free survival (LRRFS) was calculated from surgery until locoregional recurrence (LRR) (event) or the day of last follow-up without LRR (censored). Distant metastasis-free survival (DMFS) was calculated from surgery until DM/PM (event) or the day of last follow-up without DM/PM (censored). Survival was estimated by using the Kaplan–Meier method and compared with the log-rank test. All statistical analyses were performed using R software (version 3.5.0; R Project for Statistical Computing, Vienna, Austria).

**Propensity score estimation**
To evaluate differences in the baseline characteristics between groups, propensity scores were estimated with a multivariable logistic regression model in which treatment assignment was regressed on the clinicopathological variables outlined in the “Predictor variables” section. These covariates were included in the propensity score model because there was consensus that they were clinically important predictors of either treatment assignment or survival.

**Survival modeling and propensity score application**
A Cox proportional hazards model was used to compare the DFS and OS in patients treated with adjuvant CRT or ChT. To ensure balance between both groups in variables prognostic of survival and to consolidate the strength of our findings, we calculated HRs using four different methods. (1) Stratification of the propensity score was
performed by classifying patients into quintiles by their propensity scores. (2) Matching was performed by matching patients in the adjuvant ChT and CRT groups by propensity scores using variable optimal matching with a caliper of 0.1 and a ratio of 1 for the ChT and CRT groups. (3) Inverse probability of treatment weighting (IPTW) was calculated by the inverse of the propensity score of the treatment received. (4) Adjustment of the propensity score was performed by including it as a continuous covariate in the Cox regression model.

Balance diagnostics
Various graphical techniques were used to visually assess the degree of balance between treatment groups achieved with the matching approaches. These techniques included a comparison of histograms and jitter graphs depicting pre- and post-matched propensity score distributions in each treatment group (Figures S1 and S2). Standardized mean differences were computed for all the covariates in the pre- and post-matched samples (Table S1). A standardized mean difference greater than 0.1 indicated an imbalance that required further investigation.

Sensitivity analysis
Data regarding performance status, a factor associated with the choice of postoperative treatment and survival, cannot be obtained due to the retrospective nature of this study. Therefore, we conducted a sensitivity analysis with the R package “obsSens” using the approach described by Lin et al9, to investigate the potential effects of performance status on our results. We defined an Eastern Cooperative Oncology Group (ECOG) performance status of more than or equal to 2 as a bad performance status. Then, we varied the prevalence of poor performance status in the CRT and ChT groups, as well as the hazard for DFS, using estimates from prior studies. We assumed that an ECOG performance status of 2 or greater would be associated with an HR of 1.2–2.0. Using these data, we calculated adjusted HRs (with 95% CIs) for the CRT group.

Results
Study population
All the included patients were treated between January 1st, 2004, to December 31st, 2015. A detailed flow diagram of patient inclusion is shown in Figure 1A (ChT group) and Figure 1B (CRT group). There were 1787 patients identified who had undergone resection for pN3 GC at Fudan University Shanghai Cancer Center (FUSCC). We then randomly selected 800 patients to further collect their detailed clinicopathological parameters. Random selection was performed with the function “sample” in the R package “base”.

Patient characteristics
The pre-matched clinicopathological characteristics are summarized in Table 1. Patients who received adjuvant CRT were younger than those who did not. Matching by propensity scores achieved adequate balance between the adjuvant CRT and ChT groups for all covariates (Table S1, Figures S1 and S2).

Treatment delivery and safety
Patients in the CRT group received a median radiation dose of 45 Gy (range 45.0–50.4 Gy). Details of ChT regimens are summarized in Table S2. In the CRT and ChT groups, 52.6% and 49.9% of patients, respectively, completed the entire treatment regimen as planned (p=0.619). The most common grade III–IV toxicities were leucopenia/granulocytopoenia, accounting for 31.4% and 24.2% in the CRT and ChT groups, respectively. No treatment-related death occurred. Details of the toxicities are listed in Table 2.

Survival and subgroup analyses
The median follow-up for patients in the ChT group was 38 months versus 32.4 months for those in the CRT group. The addition of RT proved to be an independent prognostic factor for DFS in the multivariate analysis of the pre-matched samples (Table S3). Additionally, sex, histology, differentiation, invasion depth, nodal status, ChT courses and regimens were identified as independent prognostic factors for DFS in the Cox proportional hazard regression model selected by the Akaike Information Criterion (AIC) (Table S3).

The four different methods of propensity score analyses including adjustment, stratification, IPTW and matching yielded similar results (Table 3). In the post-matched samples, the median DFS and OS were 21.3 and 32.6 months, respectively, in the ChT group versus 23.7 and 32.9 months, respectively, in the CRT group. The 3-year DFS was 32.1% in the ChT group versus 43.4% in the CRT group (p=0.021). The 3-year OS was 45.9% in the ChT group versus 48.2% in the CRT group (p=0.159) (Table 3). Figure 2A and B shows the survival curves of
Patients who had undergone an R0 resection for stage pN3 GC (n=175)

Excluded due to unfit ChT (n=194)
- Single agent ChT (n=80)
- No detailed records of ChT (n=114)

Excluded (n=14)
- D1 lymphadenectomy (n=24)

Figure 1 Patient inclusion in the (A) adjuvant ChT group (n=365) and (B) adjuvant CRT group (n=175) (2004–2015).

Abbreviations: CRT, chemoradiotherapy; ChT, chemotherapy; FUSCC, Fudan University Shanghai Cancer Center; GC, gastric cancer; NET, neuroendocrine tumor.
Table 1  Baseline patient characteristics in the adjuvant CRT and adjuvant ChT groups in the original samples

| Variables                      | No. (%) | SMD  | Crude OR | 95% CI  | P-value |
|--------------------------------|---------|------|----------|---------|---------|
|                                | ChT Group (n=365) | CRT Group (n=175) |         |         |         |
| Sex                            |     |      |          |         |         |
| Male                           | 253 (69.3) | 116 (66.3) | 0.065 | 1.15 | 0.542 |
| Female                         | 112 (30.7) | 59 (33.7) | Ref. | 0.78–1.69 |       |
| Age (years)                    |     |      |          |         |         |
| Mean (SD)                      | 55.72 (10.69) | 52.47 (10.77) | 0.030 | 0.97 | 0.96–0.99 |
| Median (range)                 | 57 (28–80) | 54 (19–77) | Ref. |       | 0.001 |
| Location                       |     |      |          |         |         |
| GEJ+upper 1/3                  | 84 (23.0) | 31 (17.7) | 0.132 | Ref. | 0.573 |
| Middle 1/3                     | 80 (21.9) | 41 (23.4) | 1.39 | 0.79–2.43 |       |
| Lower 1/3                      | 169 (46.3) | 86 (49.1) | 1.38 | 0.85–2.27 |       |
| ≥2/3 of stomach                | 32 (8.8) | 17 (9.7) | 1.44 | 0.69–2.94 |       |
| Histology                      |     |      |          |         |         |
| Adenocarcinoma                 | 303 (83.0) | 146 (83.4) | 0.066 | Ref. | 0.921 |
| Tubular adenocarcinoma         | 17 (4.7) | 6 (3.4) | 0.73 | 0.26–1.80 |       |
| Mucinous adenocarcinoma        | 18 (4.9) | 9 (5.1) | 1.04 | 0.44–2.31 |       |
| Signet ring cell carcinoma     | 27 (7.4) | 14 (8.0) | 1.08 | 0.53–2.08 |       |
| Differentiation                |     |      |          |         |         |
| Poor                           | 238 (65.2) | 126 (72.0) | 0.147 | Ref. | 0.139 |
| Well-moderate                  | 127 (34.8) | 49 (28.0) | 0.73 | 0.49–1.08 |       |
| No. of positive LNs            |     |      |          |         |         |
| Mean (SD)                      | 13.42 (6.51) | 14.02 (6.51) | 0.091 | 1.01 | 0.99–1.04 |
| Median (range)                 | 12 (7–49) | 12 (7–38) | 0.321 |       |       |
| No. of retrieved LNs           |     |      |          |         |         |
| Mean (SD)                      | 26.52 (9.50) | 26.67 (8.81) | 0.017 | 1.00 | 0.98–1.02 |
| Median (range)                 | 24 (15–74) | 25 (15–67) | 0.855 |       |       |
| Invasion depth                 |     |      |          |         |         |
| T1                             | 12 (3.3) | 6 (3.4) | 0.134 | Ref. | 0.510 |
| T2                             | 13 (3.6) | 11 (6.3) | 1.69 | 0.48–6.27 |       |
| T3                             | 111 (30.4) | 55 (31.4) | 0.99 | 0.36–2.98 |       |
| T4                             | 229 (62.7) | 103 (58.9) | 0.90 | 0.34–2.64 |       |
| Nodal status                   |     |      |          |         |         |
| N3a                            | 269 (73.7) | 118 (67.4) | 0.138 | Ref. | 0.158 |
| N3b                            | 96 (26.3) | 57 (32.6) | 1.35 | 0.91–2.00 |       |
| AJCC stage                     |     |      |          |         |         |
| IIB                            | 12 (3.3) | 6 (3.4) | 0.134 | Ref. | 0.510 |
| IIIA                           | 13 (3.6) | 11 (6.3) | 1.69 | 0.48–6.27 |       |
| IIIB                            | 111 (30.4) | 55 (31.4) | 0.99 | 0.36–2.98 |       |
| IIIC                            | 229 (62.7) | 103 (58.9) | 0.90 | 0.34–2.64 |       |
| ChT courses                    |     |      |          |         |         |
| Complete                       | 182 (49.9) | 92 (52.6) | 0.054 | Ref. | 0.619 |
| Incomplete                     | 183 (50.1) | 83 (47.4) | 0.90 | 0.62–1.29 |       |

(Continued)
DFS and OS in the ChT and CRT groups in the post-matched dataset.

The subgroup analyses showed heterogeneity in the HRs of subgroups defined by sex, age, tumor location, differentiation, invasion depth and nodal status. However, an interaction between treatment and the subgroup of interest remained in only the subgroups defined by sex and age (Figure S3).

Patterns of recurrence

The patterns of recurrence were analyzed based on the post-matched samples. During the follow-up period, 117 and 103 patients developed recurrence in the ChT and CRT groups, respectively; of these, 18 and 20 patients in the ChT and CRT groups, respectively, could not be further classified because the sites of recurrence were unspecified, leaving 99 and 83 patients in the ChT and CRT groups, respectively, with specified recurrences for further analysis. The patterns of total recurrence are summarized in Table 4. There were more patients with isolated LRR in the ChT group than in the CRT group. With regard to the sites of recurrence, total RR was significantly decreased in the CRT group. The LRRFS in the CRT group was significantly better than that in the ChT group, but the DMFS was not different between the two groups (Figure 2C and D).

Sensitivity analysis

Sensitivity analyses including patients who were initially excluded from the analysis because either their lymphadenectomy was less than D2 or they were unfit for ChT did not alter the results (Table S4)

Assuming an HR of 1.2, the advantage of adjuvant CRT was relatively robust to the effects of poor performance status (Figure S4). However, when assuming an HR between 1.6 and 2.0, the advantage of adjuvant CRT remained only if the percentages of patients with poor performance were equal between the two groups.

Discussion

Although the risk of LRR is approximately 32.4% for patients with N3 GC, it is possible that patients with high nodal burdens, such as those with N3 disease, may not benefit from radiotherapy owing to early peritoneal or systemic failure and death.10 The current study has demonstrated that the addition of RT to adjuvant ChT can significantly improve DFS in this group of patients. These survival results remained consistent and stable in the multivariate analysis and with the four different methods based on propensity score analyses. In addition, the current study has a relatively large sample size among retrospective studies.

The possible reasons for improved DFS but not OS can be explained by the patterns of recurrence. Compared with adjuvant ChT, CRT significantly reduced the risks of LRR; however, the unimproved DMFS might have offset the LRRFS benefits from CRT. This view is also supported by several meta-analyses.11–13 Moreover, the palliative treatment after recurrence that these patients received may have influenced OS.

The results of the subgroup analyses indicated that several clinicopathological factors can be used to determine which patients will benefit from CRT. The finding that patients with well to moderately differentiated tumors benefitted more from CRT is partially consistent with the results of the ARTIST trial, which used Lauren’s classification to define subgroups.14 Another implication is that patients with N3a, but not N3b, could benefit from CRT. N3 was subclassified into N3a and N3b in the 7th edition of the AJCC staging system, but the two subclassifications do not differ with regard to the final pathologic TNM stage. However, several studies have demonstrated that the two N3 subclassifications (N3a vs N3b) have significantly different survival profiles.15–19 Thus, finally, the N3 subclassification has been incorporated into the final staging stratification in the 8th edition of the AJCC staging system. The current results are consistent with those of
Table 2 Grade III–IV toxicities

| Toxicity                           | ChT group (n=153) | CRT group (n=153) |
|------------------------------------|-------------------|-------------------|
| Leukopenia/granulocytopenia        | 37 (24.2)         | 48 (31.4)         |
| Thrombocytopenia                   | 4 (2.6)           | 5 (3.3)           |
| Anemia                             | 0 (0)             | 3 (2.0)           |
| Liver dysfunction                  | 1 (0.7)           | 2 (1.3)           |
| Anorexia                           | 7 (4.6)           | 17 (11.1)         |
| Nausea                             | 11 (7.2)          | 15 (9.8)          |
| Vomiting                           | 8 (5.2)           | 9 (5.9)           |
| Diarrhea                           | 4 (2.6)           | 3 (2.0)           |

Abbreviations: ChT, chemotherapy; CRT, chemoradiotherapy.

Table 3 Effect of adjuvant CRT versus adjuvant ChT on DFS and OS in patients with N3 gastric or gastroesophageal adenocarcinoma after D2/R0 resection

| Model                        | Sample size | DFS (95% CI)       | P-value | OS (95% CI)     | P-value |
|------------------------------|-------------|--------------------|---------|----------------|---------|
| Unadjusted                   | 175 vs 365  | 0.82 (0.65–1.02)   | 0.077   | 0.98 (0.78–1.23)| 0.860   |
| Cox regression model         | 175 vs 365  | 0.75 (0.59–0.95)   | 0.019   | 0.91 (0.72–1.17)| 0.475   |
| Propensity score-based models|             |                    |         |                |         |
| Adjusted                     | 175 vs 365  | 0.75 (0.59–0.95)   | 0.018   | 0.91 (0.72–1.16)| 0.464   |
| Stratified                   | 175 vs 365  | 0.76 (0.60–0.97)   | 0.027   | 0.93 (0.73–1.18)| 0.550   |
| Weighted (IPTW)              | 175 vs 365  | 0.80 (0.68–0.95)   | 0.012   | 0.97 (0.82–1.20)| 0.746   |
| Matched                      | 153 vs 153  | 0.72 (0.55–0.95)   | 0.021   | 0.82 (0.62–1.08)| 0.159   |

Abbreviations: ChT, chemotherapy; CRT, chemoradiotherapy; DFS, disease-free survival; IPTW, inverse probability of treatment weighting; OS, overall survival.

Table 4 Patterns of total recurrence over the entire follow-up period

|               | No. (%)      | ChT group (n=153) | CRT group (n=153) |
|---------------|--------------|-------------------|-------------------|
| Single site   |              |                   |                   |
| LRR           | 21 (13.7)    | 12 (7.8)          |
| DM            | 25 (16.3)    | 21 (13.7)         |
| PM            | 20 (13.1)    | 26 (17.0)         |
| Two sites     |              |                   |                   |
| LRR+DM        | 14 (9.2)     | 8 (5.2)           |
| LRR+PM        | 7 (4.6)      | 7 (4.6)           |
| DM+PM         | 6 (3.9)      | 9 (5.9)           |
| Three sites   |              |                   |                   |
| LRR+DM+PM     | 6 (3.9)      | 0 (0)             |
| Total recurrent sites | 141 | 110              |
| Local recurrence | 13 (9.2)   | 11 (10.0)         |
| Regional recurrence | 38 (27.0) | 18 (16.4)        |
| Peritoneal metastasis | 39 (27.7) | 43 (39.1)        |
| Distant metastasis | 51 (36.2) | 38 (34.5)        |

Abbreviations: ChT, chemotherapy; CRT, chemoradiotherapy; DM, distant metastasis; LRR, locoregional recurrence; PM, peritoneal metastasis.
previous studies. However, can patients with N3b GC be further divided? The optimal cutoff value for the number of positive LNs remains to be investigated.

The appropriate number of courses of ChT has not been determined yet. In this study, we used the number of cycles used in the ARTIST trial to dichotomize the number of courses as a variable and found that an incomplete ChT course was an independent poor prognostic factor. This finding was consistent with the results of a previous study showing that receiving insufficient courses of ChT is associated with a poor prognosis.

The optimal regimen for adjuvant ChT is also undetermined. The subgroup analyses of the ACTS-GC trial showed that the survival benefit associated with S1 treatment was limited to patients with N0-1 disease, and a single cytotoxic agent for patients with N2-3 GC is not enough. Thus, we excluded patients who received only single-agent ChT. Moreover, the selection between double- and triple-drug regimens is under debate. Both XELOX and ECF/capecitabine (ECX) have been proven to be effective in phase III clinical trials. However, ECF/ECX, which was established as a standard regimen in the MAGIC trial, has been questioned due to its poor performance, which is no better than that of single-agent ChT. Coincidentally, the triple-drug regimens used in the current study, which were mostly EOF/EOX, were identified as an independent prognostic factor for poor DFS. Recently, a new standard regimen has been established for perioperative ChT in a phase II/III clinical trial conducted in Germany; this new regimen, 5-FU + leucovorin + oxaliplatin + docetaxel (FLOT), has rather good efficacy and acceptable safety. Although, FLOT shows promise, the efficacy and safety of FLOT for Asian patients needs further validation.

As previously mentioned, the best treatment strategy for patients with N3 GC is controversial. Due to severe LN metastasis, they may be evaluated as having potentially resectable disease in practice, but surgery is high risk for these patients. Even if R0 resection is achieved, the risk of postoperative recurrence is high, and long-term survival is poor. Additionally, the completion rate of postoperative treatment was poor in our study. Nearly half of the patients in each group did not complete the planned adjuvant therapy. The same
situation has also occurred in other studies, such as the CRITICS and MAGIC trials. Compared with adjuvant therapy, neoadjuvant therapy has several advantages. This treatment modality can downstage tumor and increase the rate of R0 resection. Additionally, patients tolerate neoadjuvant therapy better than adjuvant therapy. Thus, perhaps more studies should emphasize the application of neoadjuvant therapy for N3 GC patients.

There are limitations to our study. First, as a retrospective study, different regimens of adjuvant ChT were used, and nearly half of the patients in the CRT group underwent surgery in other hospitals and received adjuvant CRT at our center. The heterogeneity of the treatments could affect the quality of this study. Second, Lauren’s typing has been shown to be a significant prognostic factor. However, it was not included in the study due to data missing. In addition, the baseline histology was not very informative due to the two versions of WHO classification adopted in the population. Thus, the category of adenocarcinoma actually included adenocarcinoma (8140/3) and mixed adenocarcinoma (8255/3). However, this information insufficiency could be complemented by information concerning the tumor differentiation. Third, we applied a 1:1 ratio to conduct the propensity score matching. Thus, the sample was, to some extent, wasted, but the matching was more accurate. As the subgroup analyses were performed based on the post-matched dataset, interpretation should be cautious, as the sample size was not large enough. Fourth, as the cases in this study were retrieved over a long time period (2004–2015), the ChT regimens changed over time. Thus, the course and regimen of ChT were also included into the propensity score analyses, although they are not, strictly speaking, baseline characteristics. Therefore, the regimens of ChT (double vs triple) between the two groups were balanced.

Conclusion
In summary, in this retrospective study, compared with adjuvant ChT, adjuvant CRT was associated with improved DFS in patients with N3 GC after D2/R0 resection. Patients with lymph node metastases should be further stratified when selecting patients for adjuvant CRT.

Ethics approval and informed consent
The study was approved by the Institutional Review Board of Fudan University Shanghai Cancer Center. Written informed consent was obtained from each patient.

Data sharing statement
Data is available upon request via email.

Acknowledgments
We appreciate Ms. Hua Sheng for her arduous work of follow-up. This study was supported by the National Natural Science Foundation of China (Grant No. 81572955).

Disclosure
The authors declare that they have no conflicts of interest in this work.

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(2):87–108. doi:10.3322/caac.21262
2. Chang JS, Lim JS, Noh SH, et al. Patterns of regional recurrence after curative D2 resection for stage III (N3) gastric cancer: implications for postoperative radiotherapy. Radiother Oncol. 2012;104(3):367–373. doi:10.1016/j.radonc.2012.08.017
3. Lee J, Lim DH, Kim S, et al. Phase III trial comparing capecitabine plus cisplatin versus capecitabine plus cisplatin with concurrent capcitabine radiotherapy in completely resected gastric cancer with D2 lymph node dissection: the ARTIST trial. J Clin Oncol. 2012;30(3):268–273. doi:10.1200/JCO.2011.39.1953
4. Kilic L, Ordu C, Ekenel M, et al. Comparison of two different adjuvant treatment modalities for pN3 gastric cancer patients after D2 lymph node dissection: can we avoid radiotherapy in a subgroup of patients? Med Oncol. 2013;30(3):660. doi:10.1007/s12032-013-0660-2
5. Fan M, Li G, Shen L, Zhang H, Liang L, Zhang Z. Identification of patients with lymph node metastasis from gastric cancer who may benefit from adjuvant chemoradiotherapy after D2 dissection-do N3 patients benefit from additional radiation? Br J Radiol. 2016;89(1059):20150758. doi:10.1259/bjr.20150758
6. Peng J, Zhou F, Dai J, et al. Benefit from adjuvant chemoradiation to resected stage IIIC gastric cancer patients with D2 lymph node dissection. J Clin Oncol. 2014;32(Suppl.):Abstract e15028. doi:10.1200/JCO.2013.54.6911
7. Nam H, Lim DH, Kim S, et al. A new suggestion for the radiation target volume after a subtotal gastrectomy in patients with stomach cancer. Int J Radiat Oncol Biol Phys. 2008;71(2):448–455. doi:10.1016/j.ijrobp.2007.09.055
8. Lee J, Lim DH, Kim S, et al. Phase III trial comparing capecitabine plus cisplatin versus capecitabine plus concurrent capcitabine radiotherapy in completely resected gastric cancer with D2 lymph node dissection. The ARTIST trial. J Clin Oncol. 2012;30(3):268–273.
9. Lin DY, Psaty BM, Kronmal RA. Assessing the sensitivity of regression results to unmeasured confounders in observational studies. Biometrics. 1998;54(3):948–963.
10. Chang JS, Kim KH, Yoon HI, et al. Locoregional relapse after gastrectomy with D2 lymphadenectomy for gastric cancer. Br J Surg. 2017;104(7):877–884. doi:10.1002/bjs.10502
11. Dai Q, Jiang L, Lin R, et al. Adjuvant chemoradiotherapy versus chemotherapy for gastric cancer: A meta-analysis of randomized controlled trials. J Surg Oncol. 2015;111(3):277–284. doi:10.1002/jso.23795
12. Huang Y, Yang Q, Zhou S, et al. Postoperative chemoradiotherapy versus postoperative chemotherapy for completely resected gastric cancer with D2 lymphadenectomy: a meta-analysis. PLoS One. 2013;8(7):e68939. doi:10.1371/journal.pone.0068939
13. Zhou ML, Kang M, Li GC, Guo XM, Zhang Z. Postoperative chemoradiotherapy versus chemotherapy for R0 resected gastric cancer with D2 lymph node dissection: an up-to-date meta-analysis. *World J Surg Oncol*. 2016;14(1):209. doi:10.1186/s12957-016-0957-7

14. Park SH, Sohn TS, Lee J, et al. Phase III trial to compare adjuvant chemotherapy with capecitabine and cisplatin versus concurrent chemoradiotherapy in gastric cancer: final report of the adjuvant chemoradiotherapy in stomach tumors trial, including survival and subset analyses. *J Clin Oncol*. 2015;33(28):3130–3136. doi:10.1200/JCO.2014.58.3930

15. Chae S, Lee A, Lee JH. The effectiveness of the new (7th) UICC N classification in the prognosis evaluation of gastric cancer patients: a comparative study between the 5th/6th and 7th UICC N classification. *Gastric Cancer*. 2011;14(2):166–171. doi:10.1007/s10120-011-0024-6

16. Chen S, Zhao BW, Li YF, et al. The prognostic value of harvested lymph nodes and the metastatic lymph node ratio for gastric cancer patients: results of a study of 1,101 patients. *PLoS One*. 2012;7(11):e49424. doi:10.1371/journal.pone.0049424

17. Jun KH, Lee JS, Kim JH, Kim JJ, Chin HM, Park SM. The rationality of N3 classification in the 7th edition of the international union against cancer TNM staging system for gastric adenocarcinomas: a case-control study. *Int J Surg*. 2014;12(9):893–896. doi:10.1016/j.ijsu.2014.06.014

18. Yeh CN, Wang SY, Hsu JT, et al. N3 subclassification incorporated into the final pathologic staging of gastric cancer: a modified system based on current AJCC staging. *Medicine (Baltimore)*. 2015;94(8):e575. doi:10.1097/MD.0000000000000874

19. Cheng Y, Huang J, Zhu Z, Xu H. [Rationality of N3 classification and its sub-classification in TNM classification system for advanced gastric cancer]. *Zhonghua Wei Chang Wai Ke Za Zhi*. 2016;19(7):749–755.

20. Li G, Zhang Z, Ma X, Zhu J, Cai G. Postoperative chemoradiotherapy combined with epirubicin-based triplet chemotherapy for locally advanced adenocarcinoma of the stomach or gastroesophageal junction. *PLoS One*. 2013;8(1):e54233. doi:10.1371/journal.pone.0055525

21. Sasako M, Sakuramoto S, Katai H, et al. Five-year outcomes of a randomized phase III trial comparing adjuvant chemotherapy with S-1 versus surgery alone in stage II or III gastric cancer. *J Clin Oncol*. 2011;29(33):4387–4393. doi:10.1200/JCO.2011.36.5908

22. Elimova E, Janjigian YY, Mulcahy M, et al. It is time to stop using epirubicin to treat any patient with gastroesophageal adenocarcinoma. *J Clin Oncol*. 2017;35(4):475–477. doi:10.1200/JCO.2016.69.7276

23. Al-Batran S, Homann N, Schmalenberg H, et al. Perioperative chemotherapy with docetaxel, oxaliplatin, and fluorouracil/leucovorin (FLOT) versus epirubicin, cisplatin, and fluorouracil or capecitabine (ECF/ECX) for resectable gastric or gastroesophageal junction (GEJ) adenocarcinoma (FLOT4-AIO): A multicenter, randomized phase 3 trial. *J Clin Oncol*. 2017;35(15_suppl):4004.

24. Al-Batran SE, Hofheinz RD, Pauligk C, et al. Histopathological regression after neoadjuvant docetaxel, oxaliplatin, fluorouracil, and leucovorin versus epirubicin, cisplatin, and fluorouracil or capecitabine in patients with resectable gastric or gastro-oesophageal junction adenocarcinoma (FLOT4-AIO): results from the phase 2 part of a multicentre, open-label, randomised phase 2/3 trial. *Lancet Oncol*. 2016;17(12):1697–1708. doi:10.1016/S1470-2045(16)30531-9
Supplementary materials

Figure S1: Histograms depicting the (A) pre- and (B) post-matched propensity score distributions in the chemotherapy (ChT) and chemoradiotherapy (CRT) groups.

Figure S2: Jitter graph depicting the pre- and post-matched propensity score distributions in the chemotherapy (ChT) and chemoradiotherapy (CRT) group.
| Nodal status | HR and 95% CI | HR | 95% CI | No. of patients |
|--------------|---------------|----|-------|----------------|
| N3a          | 0.88          | 0.63-1.23 | 203    |
| N3b          | 0.50          | 0.31-0.79 | 103    |

**Figure S3** Forest plot of the impact of adjuvant CRT versus adjuvant ChT on DFS in patient subgroups. *P* test of interaction between the treatment and subgroup of interest.

**Abbreviations:** ChT, chemotherapy; CRT, chemoradiotherapy; DFS, disease-free survival; GEJ, gastroesophageal junction.

**Figure S4** Sensitivity analysis of the HR for disease-free survival adjusted for poor performance status.

**Abbreviations:** ChT, chemotherapy; CRT, chemoradiotherapy; HR, hazard ratio.
Table S1 Standardized mean differences of baseline characteristics in the original and matched samples

| Variables          | No. (%)                          | P-value | SMD       |
|--------------------|----------------------------------|---------|-----------|
|                    | ChT Group (n=153) | CRT Group (n=153) |       | Matched sample | Original sample |
| Sex                |                      |         |           |                 |                 |
| Male               | 52 (34.0)            | 51 (33.3) | 1.000   | 0.014           | 0.065           |
| Female             | 101 (66.0)           | 102 (66.7) |         |                 |                 |
| Age (years)        |                      |         |           |                 |                 |
| Mean (SD)          | 53.95 (10.88)        | 53.47 (10.22) | 0.699   | 0.046           | 0.303           |
| Median (range)     | 56 (28–80)           | 55 (19–77) |         |                 |                 |
| Location           |                      |         |           |                 |                 |
| GEJ+upper 1/3      | 70 (45.8)            | 74 (48.4) | 0.936   | 0.074           | 0.132           |
| Middle 1/3         | 37 (24.2)            | 36 (23.5) |         |                 |                 |
| Lower 1/3          | 33 (21.6)            | 29 (19.0) |         |                 |                 |
| ≥2/3 of stomach    | 13 (8.5)             | 14 (9.2) |         |                 |                 |
| Histology          |                      |         |           |                 |                 |
| Adenocarcinoma     | 125 (81.7)           | 126 (82.4) | 0.925   | 0.078           | 0.066           |
| Tubular adenocarcinoma | 7 (4.6)         | 6 (3.9)   |         |                 |                 |
| Mucinous adenocarcinoma | 10 (6.5)      | 8 (5.2)    |         |                 |                 |
| Signet ring cell carcinoma | 11 (7.2) | 13 (8.5) |         |                 |                 |
| Differentiation    |                      |         |           |                 |                 |
| Poor               | 107 (69.9)           | 109 (71.2) | 0.900   | 0.029           | 0.147           |
| Well-moderate      | 46 (30.1)            | 44 (28.8) |         |                 |                 |
| No. of positive LNs|                      |         |           |                 |                 |
| Mean (SD)          | 14.14 (7.39)         | 13.60 (6.51) | 0.496   | 0.078           | 0.091           |
| Median (range)     | 12 (7–49)            | 12 (7–38) |         |                 |                 |
| No. of retrieve LNs|                      |         |           |                 |                 |
| Mean (SD)          | 26.07 (9.13)         | 26.82 (9.06) | 0.474   | 0.082           | 0.017           |
| Median (range)     | 24 (15–61)           | 25 (15–67) |         |                 |                 |
| Invasion depth     |                      |         |           |                 |                 |
| T1                 | 4 (2.6)              | 6 (3.9) | 0.912   | 0.083           | 0.134           |
| T2                 | 9 (5.9)              | 9 (5.9) |         |                 |                 |
| T3                 | 46 (30.1)            | 48 (31.4) |         |                 |                 |
| T4                 | 94 (61.4)            | 90 (58.8) |         |                 |                 |
| Nodal status       |                      |         |           |                 |                 |
| N3a                | 109 (71.2)           | 111 (72.5) | 0.899   | 0.029           | 0.138           |
| N3b                | 44 (28.8)            | 42 (27.5) |         |                 |                 |
| AJCC stage         |                      |         |           |                 |                 |
| IIB                | 4 (2.6)              | 6 (3.9) | 0.912   | 0.083           | 0.134           |
| IIIA               | 9 (5.9)              | 9 (5.9) |         |                 |                 |
| IIIB               | 46 (30.1)            | 48 (31.4) |         |                 |                 |
| IIIC               | 94 (61.4)            | 90 (58.8) |         |                 |                 |
| ChT courses        |                      |         |           |                 |                 |
| Complete           | 80 (52.3)            | 79 (51.6) | 1.000   | 0.013           | 0.054           |
| Incomplete         | 73 (47.7)            | 74 (48.4) |         |                 |                 |

(Continued)
Table S1 (Continued).

| Variables                  | No. (%) | P-value | SMD | Matched sample | Original sample |
|----------------------------|---------|---------|-----|----------------|-----------------|
|                            | ChT Group (n=153) | CRT Group (n=153) |     |                |                 |
| ChT regimens               |         |         |     |                |                 |
| Double drugs               | 90 (58.8) | 90 (58.8) | 1.000 | <0.001         | 0.617           |
| Triple drugs               | 63 (41.2) | 63 (41.2) |      |                |                 |

Abbreviations: AJCC, American Joint Committee on Cancer; ChT, chemotherapy; CRT, chemoradiotherapy; GEJ, gastroesophageal junction; LNs, lymph nodes.

Table S2 ChT regimens

| Regimens                         | n (%) | ChT group (n=365) | CRT group (n=175) |
|----------------------------------|-------|-------------------|-------------------|
| Epirubicin+platinum+fluorouracil | 63 (17.3) | 83 (47.4)          |
| Taxanes+platinum+fluorouracil    | 12 (3.3)  | 2 (1.1)            |
| Platinum+fluorouracil            | 266 (72.9) | 86 (49.1)          |
| Taxanes+platinum                 | 22 (6.0)   | 3 (1.7)            |
| Others                           | 2 (0.5)   | 1 (0.6)            |

Abbreviations: ChT, chemotherapy; CRT, chemoradiotherapy.

Table S3 Univariate and multivariate analyses of prognostic factors for DFS

| Variables                  | Univariate analysis | Multivariate analysis |
|----------------------------|---------------------|-----------------------|
|                            | N       | No. of events | 3-yr DFS (%) | HR | 95% CI      | P-value | HR | 95% CI      | P-value |
| Sex                        |         |              |              |    |             |         |    |             |         |
| Male                       | 369     | 242          | 40.9%        | Ref.|               |         |    |               |         |
| Female                     | 171     | 124          | 32.5%        | 1.19| 0.96–1.48   | 0.111  | 1.28| 1.02–1.60   | 0.033  |
| Age (years)                |         |              |              |    |             |         |    |             |         |
| ≤60                        | 374     | 251          | 38.2%        | Ref.|               |         |    |               |         |
| >60                        | 166     | 115          | 38.1%        | 1.03| 0.82–1.28   | 0.804  | –  | –             | –       |
| Location                   |         |              |              |    |             |         |    |             |         |
| GEJ+upper 1/3              | 115     | 81           | 33.8%        | Ref.|               |         |    |               |         |
| Middle 1/3                 | 121     | 81           | 39.4%        | 0.90| 0.66–1.22   | 0.487  | –  | –             | –       |
| Lower 1/3                  | 255     | 169          | 40.0%        | 0.91| 0.69–1.18   | 0.461  | –  | –             | –       |
| Multiple                   | 49      | 35           | 33.9%        | 1.19| 0.80–1.77   | 0.385  | –  | –             | –       |
| Histology                  |         |              |              |    |             |         |    |             |         |
| Adenocarcinoma             | 449     | 297          | 39.9%        | Ref.|               |         |    |               |         |
| Tubular adenocarcinoma     | 23      | 17           | 20.4%        | 1.52| 0.93–2.47   | 0.095  | 1.77| 1.06–2.94   | 0.028  |
| Mucinous adenocarcinoma    | 27      | 18           | 34.3%        | 1.06| 0.66–1.70   | 0.815  | 1.01| 0.62–1.63   | 0.982  |
| Signet ring cell carcinoma | 41      | 34           | 25.2%        | 1.51| 1.06–2.16   | 0.023  | 1.35| 0.94–1.94   | 0.105  |
| Differentiation            |         |              |              |    |             |         |    |             |         |
| Poor                       | 364     | 260          | 35.2%        | Ref.|               |         |    |               |         |
| Well-moderate              | 176     | 106          | 44.2%        | 0.75| 0.60–0.94   | 0.014  | 0.78| 0.61–0.99   | 0.044  |
| Invasion depth             |         |              |              |    |             |         |    |             |         |
| T1                         | 18      | 6            | 77.8%        | Ref.|               |         |    |               |         |
| T2                         | 24      | 13           | 38.7%        | 2.04| 0.77–5.36   | 0.149  | 2.38| 0.90–6.28   | 0.079  |

(Continued)
Table S3 (Continued).

|                | Univariate analysis | Multivariate analysis |
|----------------|--------------------|-----------------------|
|                |                    |                       |
| T3             | 166                | 97                    |
|                | 45.9%              | 2.18                  |
|                | 0.96–4.98          | 0.064                 |
| T4             | 332                | 250                   |
|                | 32.2%              | 3.22                  |
|                | 1.43–7.24          | 0.005                 |
| Nodal status   |                    |                       |
| N3a            | 387                | 250                   |
|                | 41.4%              | Ref.                  |
| N3b            | 153                | 116                   |
|                | 29.7%              | 1.48                  |
|                | 1.18–1.84          | 0.001                 |
| ChT course     |                    |                       |
| Complete       | 274                | 180                   |
|                | 42.8%              | Ref.                  |
| Incomplete     | 266                | 186                   |
|                | 33.7%              | 1.29                  |
|                | 1.05–1.59          | 0.014                 |
| ChT regimens   |                    |                       |
| Double drugs   | 380                | 243                   |
|                | 41.4%              | Ref.                  |
| Triple drugs   | 160                | 123                   |
|                | 30.4%              | 1.26                  |
|                | 1.01–1.56          | 0.039                 |
| Group          |                    |                       |
| ChT            | 365                | 259                   |
|                | 36.1%              | Ref.                  |
| CRT            | 175                | 107                   |
|                | 42.5%              | 0.82                  |
|                | 0.65–1.02          | 0.077                 |

Abbreviations: ChT, chemotherapy; CRT, chemoradiotherapy; GEJ, gastroesophageal junction; LNs, lymph nodes; Ref., reference.

Table S4 Sensitivity analyses examining the effect of adjuvant CRT versus adjuvant ChT on DFS and OS in patients with N3 gastric or gastroesophageal adenocarcinoma after D2/R0 resection

| Model                              | Sample size | DFS          | OS           |
|------------------------------------|-------------|--------------|--------------|
|                                    |             | HR (95% CI) | P-value      | HR (95% CI) | P-value |
| Including patients unfit for chemotherapy |             |             |              |             |
| Unadjusted                         | 532 vs 202  | 0.78 (0.64–0.95) | 0.013 | 0.88 (0.72–1.08) | 0.211    |
| Including patients do not reach D2 lymphadenectomy | | |
| Unadjusted                         | 389 vs 193  | 0.83 (0.67–1.03) | 0.092 | 0.99 (0.80–1.23) | 0.920    |
| Propensity score adjusted          | 389 vs 193  | 0.76 (0.61–0.95) | 0.017 | 0.91 (0.73–1.10) | 0.439    |

Abbreviations: ChT, chemotherapy; CRT, chemoradiotherapy; DFS, disease-free survival; IPTW, inverse probability of treatment weighting; OS, overall survival.