Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Long-term outcomes and pretreatment immune profiling of perioperative toripalimab plus FLOT in gastric or gastroesophageal junction adenocarcinoma: a phase II trial

Hongli Li

lihongli@tjmuch.com

Tianjin Medical University Cancer Institute and Hospital

Shao-Hua Ge

Tianjin Medical University Cancer Institute and Hospital

Fenglin Zang

Tianjin Medical University Cancer Institute and Hospital

Le Zhang

Tianjin Medical University Cancer Institute and Hospital

Yuchong Yang

First Hospital of China Medical University

Jing-Jing Duan

Tianjin Medical University Cancer Institute and Hospital

Peng Ren

Tianjin Medical University Cancer Institute and Hospital

Fuliang Cao

Tianjin Medical University Cancer Institute and Hospital

Ming Bai

Tianjin Medical University Cancer Institute and Hospital

Tao Ning

Laboratory of Cancer Cell Biology, Tianjin Cancer Institute, Tianjin Medical University Cancer Institute and Hospital

Likun Zhou

Nanjing University

Feixue Wang

Tianjin Medical University Cancer Institute and Hospital

Jiayu Zhang

Tianjin Medical University Cancer Institute and Hospital

Yansha Sun

State Key Laboratory of Oncogenes and Related Genes, Shanghai Cancer Institute, Renji Hospital, Shanghai Jiaotong University School of Medicine, Shanghai

Rui Liu
Tianjin Medical University Cancer Institute and Hospital

Yi Ba
Tianjin Medical University Cancer Institute and Hospital https://orcid.org/0000-0003-0255-4376

Li-Yan Zhou
Tianjin Medical University Cancer Institute and Hospital

Xiao-Qing Li
Tianjin Medical University Cancer Institute and Hospital

Yue Song
Tianjin Medical University Cancer Institute and Hospital

Si Yiran
Tianjin Medical University Cancer Institute & Hospital

Yi-Lin Deng
Tianjin Medical University Cancer Institute and Hospital

Jingyu Deng
Tianjin Medical University Cancer Institute and Hospital https://orcid.org/0000-0003-4299-1327

Ting Deng
Tianjin Medical University Cancer Institute and Hospital https://orcid.org/0000-0003-4552-8039

Article

Keywords: gastric or gastroesophageal junction adenocarcinoma, toripalimab, FLOT, tumor immune microenvironment, stromal CD3 positivity

Posted Date: November 19th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11665269/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Long-term outcomes and pretreatment immune profiling of perioperative toripalimab plus FLOT in gastric or gastroesophageal junction adenocarcinoma: a phase II trial

Hong-Li Li [1,†,*], Shao-Hua Ge [2,†], Feng-Lin Zang [3,†], Le Zhang [2], Yu-Chong Yang [2], Jing-Jing Duan [2], Peng Ren [4], Fu-Liang Cao [5], Ming Bai [2], Tao Ning [2], Li-Kun Zhou [2], Fei-Xue Wang [2], Jia-Yu Zhang [2], Yan-Sha Sun [2], Rui Liu [2], Yi Ba [6], Li-Yan Zhou [7], Xiao-Qing Li [7], Yue Song [7], Yi-Ran Si [7], Yi-Lin Deng [7], Jing-Yu Deng [8,*], Ting Deng [2,*]

[1] Department of Gastrointestinal Medical Oncology and Phase I Clinical Trial Ward, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, State Key Laboratory of Druggability Evaluation and Systematic Translational Medicine, Tianjin Key Laboratory of Digestive Cancer, Tianjin Clinical Research Center for Cancer, Tianjin 300060, China.

[2] Department of Gastrointestinal Medical Oncology, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Tianjin 300060, China.

[3] Department of Pathology, Tianjin Medical University Cancer Institute and Hospital, Key Laboratory of Cancer Prevention and Therapy, Tianjin Clinical Research Center for Cancer, National Clinical Research Center for Cancer, Tianjin 300060, China.

[4] Department of Esophageal Cancer, Tianjin Medical University Cancer Institute & Hospital, National Clinical Research Center for Cancer, Tianjin, 300060, China.

[5] Department of Endoscopy Diagnosis and Therapy, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Tianjin Clinical Research Center for Cancer, Tianjin 300060, China.

[6] Department of Gastrointestinal Medical Oncology, Tianjin Medical University Cancer Institute & Hospital; National Clinical Research Center for Cancer, Tianjin,300060, China. Department of Medical Oncology, Department of Cancer Center, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences, Beijing, 100032, China.

[7] Phase I Clinical Trial Ward, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, State Key Laboratory of Druggability Evaluation and Systematic Translational Medicine, Tianjin 300060, China.

[8] Department of Gastric Surgery, Tianjin Medical University Cancer Institute & Hospital; National Clinical Research Center for Cancer, Tianjin 300060, China. Department of pancreatic and Gastric Surgery, National Cancer Center/ National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China.

[†] Hong-Li Li, Shao-Hua Ge and Feng-Lin Zang contributed equally to this work and share first authorship.

[*] Correspondence to: Professor Hong-Li Li, Department of Gastrointestinal Medical Oncology and Phase I Clinical

Trial Ward, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, State Key Laboratory of Druggability Evaluation and Systematic Translational Medicine, Tianjin Key Laboratory of Digestive Cancer, Tianjin Clinical Research Center for Cancer, Tianjin 300060, China. Tel: +86 18622221233. E-mail: lihongli@tjmuch.com. Or Professor Jing-Yu Deng, Department of pancreatic and Gastric Surgery, National Cancer Center/ National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China. Tel: +86 18622228656. E-mail: dengery@126.com. Or Professor Ting Deng, Department of Gastrointestinal Medical Oncology, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Tianjin 300060, China. Tel: +86 15802243063. E-mail: xymcdengting@126.com.

Abstract

Long-term outcomes after perioperative immunochemotherapy for resectable gastric or gastroesophageal junction (G/GEJ) adenocarcinoma remain incompletely characterised. In this single-arm phase II trial (NCT04354662), 51 patients with locally advanced G/GEJ received toripalimab plus fluorouracil, leucovorin, oxaliplatin and docetaxel before surgery with postoperative treatment were enrolled. The reported co-primary endpoints were primary-tumour pathological complete response and 3-year DFS. With median follow-up of 63.1 months, primary-tumour complete response was 19.6% (10/51), whereas nodal-inclusive ypT0N0 response was 13.7% (7/51). Three-year DFS was 52.9% (95% CI, 38.5%-65.5%), 5-year OS was 59.0% (42.9%-71.9%). Grade ≥3 reported treatment-related adverse events occurred in 72.5% of patients. Pretreatment biopsies from 24 patients underwent multiplex immunofluorescence. Higher stromal CD3 positivity was nominally associated with longer survival after median dichotomization, but continuous-variable associations were nonsignificant and no feature survived false-discovery-rate correction. Findings support independent investigation of stromal CD3 positivity but establish neither comparative benefit nor biomarker validity.

Trial registration: ClinicalTrials.gov (NCT04354662), first registered on 15 April 2020 and first posted on 21 April 2020.

Keywords: gastric or gastroesophageal junction adenocarcinoma; toripalimab; FLOT; tumor immune microenvironment; stromal CD3 positivity

INTRODUCTION

Gastric cancer (GC) remains a major cause of cancer-related mortality worldwide, and many patients present with locally advanced disease and a substantial risk of recurrence[1]. In China, approximately 342,000 new cases and 249,000 deaths were estimated in 2024[2]. For patients with locally advanced, resectable gastric or gastroesophageal junction (G/GEJ) adenocarcinoma, curative-intent gastrectomy with adequate lymphadenectomy remains central to treatment, whereas perioperative systemic therapy aims to eradicate occult micrometastatic disease[3]. Strategies combining immune checkpoint inhibitors (ICIs) with chemotherapy, with or without antiangiogenic agents, have been investigated to improve outcomes[4]. Fluorouracil, leucovorin, oxaliplatin and docetaxel (FLOT) is an established perioperative chemotherapy backbone following the survival benefit shown in the FLOT4 trial[5]. Nevertheless, recurrence remains a major challenge, creating a need for more effective perioperative strategies.

Randomized trials have clarified the role of perioperative immunotherapy. The DANTE trial reported greater pathological regression with atezolizumab plus FLOT, particularly in biologically enriched subgroups[6]. KEYNOTE-585 increased pathological complete response (pCR) rates with pembrolizumab plus chemotherapy, although the prespecified statistical boundary for event-free survival (EFS) was not crossed at the interim analysis and the final analysis did not establish an overall survival (OS) benefit[7]. MATTERHORN, in contrast, improved EFS with durvalumab plus FLOT[8]. These findings support the clinical potential of perioperative chemoimmunotherapy while highlighting differences in efficacy across treatment regimens, patient populations and biological subtypes.

Toripalimab-based perioperative strategies have also shown promising activity. The randomized NEOSUMMIT-01 trial demonstrated improved pathological response with toripalimab plus oxaliplatin-based chemotherapy, with subsequent follow-up reporting benefits in 3-year EFS and OS[9]. More recently, Zhao and colleagues reported a phase II trial of neoadjuvant toripalimab plus capecitabine and oxaliplatin (CapeOX) in Epstein-Barr virus (EBV)-associated G/GEJ adenocarcinoma in Nature Communications[10]. Among 16 resected

patients, major pathological response (MPR) and nodal-inclusive pCR (ypT0N0) rates were 37.5% and 25.0%, respectively. Multiplex immunofluorescence analyses further linked higher pretreatment CD8-positive T-cell infiltration with pathological response and characterized treatment-associated changes in the tumor immune microenvironment[10]. However, these findings were derived from an EBV-selected population receiving a CapeOX-based regimen. Meta-analytic evidence also supports perioperative chemoimmunotherapy while highlighting heterogeneity across studies[11]. Mature survival outcomes and the prognostic relevance of pretreatment immune features after toripalimab combined specifically with FLOT remain less well characterized.

Against this background, we report long-term follow-up from an investigator-initiated, single-arm phase II trial of perioperative toripalimab plus FLOT in 51 patients with locally advanced, resectable G/GEJ adenocarcinoma. We evaluated pathological regression, disease-free survival (DFS), OS and safety, together with exploratory tumor immune microenvironment (TIME) profiling of 36 compartment-specific immune features in pretreatment biopsies from 24 patients. By integrating mature clinical outcomes with tumor-parenchymal and stromal immune measurements, this study aimed to provide complementary clinical evidence and identify candidate prognostic features for further validation. Given the single-arm design, comparative treatment effects could not be estimated, and biomarker findings were considered hypothesis-generating.

Results

Patient disposition and baseline characteristics

As illustrated in Figure 1, 55 patients were screened, 4 did not meet the eligibility criteria. The 51 patients who initiated perioperative toripalimab plus FLOT comprised the full clinical and safety analysis sets. All 51 underwent gastrectomy and were included in the clinical, survival and overall safety analyses. At the data cutoff of 7 July 2026, 36 patients had completed planned perioperative treatment, 6 had discontinued or modified treatment because of adverse events or intolerance, 2 had stopped treatment at their own request or withdrawn, and 7 had an

unknown reason. No patient remained on active protocol treatment.

Baseline characteristics were shown in Table 1, median age was 59 years (IQR: 55.0-65.5, range: 28-75 years). ECOG performance status was 0 in 25 patients (49.0%) and 1 in 26 patients (51.0%). The primary tumor was located in the stomach in 45 patients (88.2%), at the gastroesophageal junction in 2 (3.9%). The study population predominantly had locally advanced disease: 48 patients (94.1%) had cT4 tumors, 44 (86.3%) had clinically node-positive disease, and 41 (80.4%) had stage III disease. Regarding baseline molecular characteristics, 45 patients (88.2%) had proficient mismatch repair (pMMR) tumors and 4 (7.8%) had deficient mismatch repair (dMMR) tumors. Epstein-Barr virus-encoded RNA (EBER) positivity was detected in 6 patients (11.8%). Human epidermal growth factor receptor 2 (HER2) immunohistochemistry was scored as 0 in 31 patients (60.8%), 1+ in 5 (9.8%), and 2+ in 11 (21.6%).

Therapeutic outcomes and pathological response

Among the 9 RECIST-evaluable patients, 5 had a partial response and 4 had stable disease, no complete response or progressive disease was recorded. The ORR was therefore 55.6% (95%CI: 21.2%-86.3%) and the DCR was 100% (95%CI: 66.4%-100.0%). As shown in Supplementary Table S1, all 51 patients underwent gastrectomy after neoadjuvant therapy. 19 (37.3%) underwent total gastrectomy, 16 (31.4%) distal gastrectomy, 14 (27.5%) proximal gastrectomy and 2 (3.9%) combined proximal gastrectomy with esophageal resection. D2 lymphadenectomy was recorded in 42 patients (82.4%), D1+ in 6 and another extent in 3. R0 resection was documented in 43 patients (84.3%). The median number of examined nodes was 28 (IQR: 18-47, range: 7-81). The median number of positive nodes was 1 (IQR: 0-3.5, range: 0-22), and 19 patients (37.3%) had no positive nodes.

Pathological tumor regression was evaluable in 50 of the 51 treated patients. Among these 50 patients, TRG0, TRG1, TRG2 and TRG3 were observed in 10 (20.0%), 6 (12.0%), 20 (40.0%) and 14 (28.0%) patients, respectively. For protocol-defined binary endpoints, all 51 treated patients were retained in the denominator. As illustrated in Figure 2A, the protocol-defined pCR rate (TRG0) was 19.6% (10/51, 95%CI: 9.8%-33.1%), and the MPR rate (TRG0-

1) was 31.4% (16/51, 95%CI: 19.1%-45.9%). A stringent nodal-inclusive pCR, defined as ypT0N0, was achieved in 7 patients (13.7%, 95%CI: 5.7%-26.3%). 3 patients with complete regression of the primary tumor (TRG0) had residual metastatic disease in regional lymph nodes and therefore did not meet the ypT0N0 definition. Exploratory subgroup analyses of protocol-defined pCRs were presented in Supplementary Table S2. The pCR rate was higher in patients with dMMR tumors than in those with pMMR tumors (75.0% [3/4] versus 11.1% [5/45], Fisher's exact *P*=0.011). No clear differences in pCR were observed across the other clinical subgroups. Given the small number of patients with dMMR tumors, this finding should be interpreted as exploratory.

Perioperative treatment exposure and long-term survival outcomes

At the data cutoff (7 July 2026), the reverse-Kaplan-Meier median follow-up duration was 63.1 months (range: 6.9-86.1 months, 95%CI: 52.2-67.5 months). Patient-level perioperative treatment exposure and longitudinal clinical outcomes were shown in Figure 2B. All 51 patients underwent surgery after neoadjuvant treatment, with substantial interpatient variability in postoperative treatment exposure and follow-up duration. During follow-up, 24 patients experienced a DFS event and 19 patients died. As shown in Figure 3, the median OS of the 51 patients was not reached, the lower 95% confidence limit was 50.1 months, and 3-year and 5-year OS rates were 69.8% (95% CI: 54.9%-80.6%) and 59.0% (42.9%-71.9%), respectively. Median DFS was not reached, with a lower 95% confidence limit of 14.1 months. Both 3-year and 5-year DFS rates were 52.9% (38.5%-65.5%).

Exploratory univariable subgroup analyses of DFS and OS were shown in Supplementary Table S3. Patients who achieved pCR had significantly longer DFS and OS than those who did not, no DFS event or death occurred among the 10 patients with pCR during follow-up (log-rank *P*=0.003 for DFS and *P*=0.009 for OS, Supplementary Figure S1). MPR and R0 resection were also associated with more favorable long-term outcomes. In parsimonious multivariable Cox models including MPR and R0 resection, MPR remained associated with longer DFS (adjusted HR=0.127, 95% CI: 0.029-0.547, *P*=0.006) and OS (adjusted HR=0.126, 95% CI: 0.028-0.573, *P*=0.007). R0 resection was likewise associated with longer DFS (adjusted

HR=0.293, 95% CI: 0.113-0.761, *P*=0.012) and OS (adjusted HR=0.156, 95% CI: 0.055-0.444, *P*<0.001) (Supplementary Table S4).

Safety profile of toripalimab plus FLOT

As shown in Table 2, all 51 patients who received at least one dose of neoadjuvant toripalimab plus FLOT experienced at least one TRAE, corresponding to an incidence of 100.0% (95% CI: 93.0%-100.0%). Grade ≥3 TRAEs occurred in 37 patients (72.5%, 95% CI: 58.3%-84.1%), including grade 4 events in 10 patients (19.6%); no grade 5 TRAEs were recorded. The most common TRAEs of any grade were neutropenia (90.2%), leukopenia (86.3%), nausea (74.5%), anemia (72.5%), fatigue (62.7%), alopecia (62.7%), decreased appetite (54.9%) and dizziness (43.1%). The most frequent grade ≥3 TRAEs were neutropenia (51.0%) and leukopenia (27.5%). Other grade ≥3 TRAEs were less common, including anemia, increased alanine/aspartate aminotransferase levels and rash (5.9% each), oral mucositis and hypokalemia (3.9% each), and hypoalbuminemia, increased bilirubin and bone marrow suppression (2.0% each).

Events of potential immune-mediated clinical interest included rash in 8 patients (15.7%), hypothyroidism in 3 (5.9%), hyperthyroidism in 1 (2.0%) and pneumonitis in 1 (2.0%). No treatment-related deaths were reported.

Pretreatment TIME landscape

Of the 51 treated patients, 24 had quality-controlled TIME data (9 TRG0-1, 15 TRG2-3) and were included in the translational analysis, the remaining 27 comprised the non-TIME cohort. As shown in Supplementary Table S5, baseline characteristics were generally comparable between the groups. No significant differences were observed in age, primary tumor site, clinical T or N stage, overall clinical stage, HER2 expression or EBER status. However, the TIME cohort had a numerically higher proportion of male patients (87.5% vs 63.0%, *P*=0.058). Notably, no known dMMR tumors were represented in the TIME cohort, whereas 4 of 27 patients (14.8%) in the non-TIME cohort had dMMR tumors, the overall distribution of mismatch repair status showed a borderline difference between the groups (*P*=0.055).

Multiplex immune profiling yielded 36 prespecified raw features comprising the densities

and positive-cell proportions of CD3, CD8, PD-1, PD-L1, CD56, CD8+PD-1+, CD3+PD-1+ and CD68/CD163-defined macrophage populations in the tumor and stromal compartments. As shown in Figure 4A, pretreatment TIME profiling was available for 24 patients, including 9 with favorable pathological regression (TRG0-1) and 15 with TRG2-3. The heatmap of 36 immune features showed substantial interpatient heterogeneity, including marked variation among patients within the same TRG category. None of the 36 individual raw TIME features differed significantly between the TRG0-1 and TRG2-3 groups after correction for multiple testing, indicating that no single dominant pretreatment immune feature characterized pathological response in this exploratory cohort.

Interestingly, as shown in Figure 4B, stromal CD3 positivity was numerically higher in TRG0-1 than in TRG2-3 (median 31.9% versus 20.3%, *P*=0.180), but the difference was not statistically significant and was not interpreted as a response biomarker. Within the TIME subset, 10 DFS events and 7 deaths were recorded, and reverse Kaplan-Meier median follow-up was 65.7 months. The 5-year DFS and OS estimates were 58.3% and 67.4%, respectively. TRG0-1 patients experienced one DFS event and one death, whereas TRG2-3 patients experienced 9 DFS events and 6 deaths; the corresponding exploratory log-rank P values were 0.035 for DFS and 0.078 for OS, respectively.

Stromal CD3 positivity and long-term outcomes

Representative multiplex immunofluorescence images from two patients in the stromal CD3-high subgroup were shown in Figure 5. The composite and corresponding CD3 single-channel images illustrated CD3-positive cells within the stromal compartment surrounding tumor epithelial structures. These images were illustrative, quantitative stromal CD3 positivity was derived from digital image analysis across the 24-patient TIME cohort rather than from visual scoring of the representative fields. Cohort-median grouping yielded 12 patients in each group. Median OS was not reached in the high group versus 59.2 months in the low group, and 5-year OS was 88.9% versus 45.0% (log-rank *P*=0.023, Figure 4C). Median DFS was not reached versus 15.1 months, and 5-year DFS was 83.3% versus 33.3% (*P*=0.010; Figure 4D). Across identical median-split analyses of all 36 features, stromal CD3 positivity was the only feature

nominally associated with both DFS (HR=0.16, 95% CI: 0.03-0.78, log-rank *P*=0.010, q=0.367) and OS (HR=0.13, 95% CI: 0.02-1.05, log-rank *P*=0.023, q=0.822). No feature met q<0.05 (Supplementary Figures S2 and S3).

Stromal CD3 positivity was also analyzed as a continuous variable after log1p transformation and standardization. In univariable Cox models, higher stromal CD3 positivity showed directionally favorable but nonsignificant associations with DFS (HR=0.63, 95% CI: 0.36-1.09, *P*=0.095) and OS (HR=0.72, 95% CI: 0.37-1.41, *P*=0.342). After adjustment for TRG category, dichotomized CD3-high status showed nominal associations with improved DFS (HR=0.21, 95% CI: 0.03-0.88, *P*=0.032) and OS (HR=0.15, 95% CI: 0.01-0.88, *P*=0.034), whereas the corresponding continuous-variable models remained statistically nonsignificant (Supplementary Table S6).

Discussion

This phase II study combined long-term clinical follow-up with pretreatment TIME profiling after perioperative toripalimab plus FLOT. In the 51-patient cohort, primary-tumor pCR was 19.6%, nodal-inclusive ypT0N0 response was 13.7%, and estimated 5-year OS and DFS were 59.0% and 52.9%, respectively. Most patients had cT4 and clinically node-positive disease. These outcomes provide descriptive information on a high-risk cohort treated before definitive perioperative immunotherapy evidence became available. They complement randomized trials but do not establish superiority over chemotherapy alone. The translational analysis identified stromal CD3 positivity as a candidate prognostic signal, but no marker survived panel-wide FDR correction. Long-term follow-up and biological profiling were therefore the principal contributions, with distinct limitations applying to the clinical and translational components.

FLOT regimen had randomized efficacy evidence and supportive real-world data[5,12]. Other single-arm perioperative immunotherapy studies had integrated pathological outcomes with multi-omics analyses, illustrating the potential value of linking clinical and biological observations[13]. However, pathological improvement did not necessarily establish survival benefit. KEYNOTE-585 increased pCR without meeting its interim EFS boundary, and the final

analysis did not establish an OS advantage[7,14]. DANTE reported greater pathological regression with atezolizumab plus FLOT, whereas MATTERHORN improved EFS and increased pCR from 7.2% to 19.2% with durvalumab plus FLOT[6,8]. The final OS analysis of MATTERHORN, published in September 2026, reported a significant survival benefit with perioperative durvalumab plus FLOT compared with FLOT alone (HR=0.78, 95% CI: 0.63-0.96, P=0.021)[15]. This randomized evidence provides contemporary context for interpreting our study. Although our cohort provides longer-term descriptive outcomes with toripalimab plus FLOT, it cannot establish comparative efficacy or therapeutic equivalence to durvalumab-based perioperative treatment. DRAGON IV also increased pCR with camrelizumab, low-dose rivoceranib and chemotherapy[4]. Differences among trial populations, chemotherapy backbones, schedules and statistical hypotheses complicate cross-trial interpretation[16]. Accordingly, efficacy estimates for one ICI regimen should not be transferred directly to another, and our study could not isolate the incremental effect of toripalimab.

Evidence specifically involving toripalimab is also relevant. In NEOSUMMIT-01, adding toripalimab to SOX or XELOX increased TRG0-1 response from 20.4% to 44.4% and ypT0N0 response from 7.4% to 22.2%[17]. Three-year EFS was 74.7% versus 56.2% (HR=0.51), and OS was 81.3% versus 72.2% (HR=0.45) [9]. A meta-analysis of 7 randomized trials involving 2510 patients estimated pCR rates of 17.6% with chemoimmunotherapy versus 6.1% with chemotherapy and pooled EFS and OS HRs of 0.76 and 0.82, respectively. Broader syntheses also favored perioperative chemoimmunotherapy while emphasizing heterogeneity[11,18]. Our FLOT-based cohort adds longer descriptive follow-up with a different chemotherapy backbone. Its 5-year estimates should not be compared directly with 3-year results from randomized or molecularly enriched cohorts. Differences in clinical stage, molecular composition, surgical selection and treatment completion may account for substantial variation independently of the checkpoint inhibitor. A recent phase II study by Zhao et al. provided an informative comparison. It investigated neoadjuvant toripalimab plus CapeOX specifically in Epstein-Barr virus-associated G/GEJ[10]. Among 16 patients undergoing surgery, the investigators reported an MPR rate of 37.5% and a nodal-inclusive pCR rate of 25.0%, compared with 31.4% and 13.7%,

respectively, in our cohort. These numerical differences should not be interpreted as evidence of differential regimen efficacy. Zhao et al. enrolled exclusively EBV-positive, pMMR tumors, whereas our population was not selected by EBV status and included only 6 EBER-positive patients. Furthermore, different chemotherapy backbones, pathological assessment populations and response definitions complicate cross-trial interpretation. Median follow-up was also shorter in the EBV-selected study (29.8 months), limiting comparisons of long-term clinical outcomes. Together, these findings highlight the importance of molecular context when interpreting pathological responses to perioperative immunochemotherapy. Our study provides complementary evidence through longer-term survival observations in a clinically high-risk, non-EBV-selected population treated with toripalimab plus FLOT, without establishing the relative efficacy of either regimen.

PANDA linked pathological response with serial immune profiling, illustrating the value of integrating clinical and biological data [19]. However, pathological grading systems are not interchangeable[20]. Our primary-tumor TRG0 endpoint differed from the nodal-inclusive ypT0N0 response: 3 patients had complete regression in the primary tumor bed but residual regional nodal disease. This distinction is important because ypT0 does not necessarily imply eradication of regional metastases[21]. Reporting both endpoints avoids treating the 19.6% primary-tumor response rate as equivalent to the 13.7% ypT0N0 rate. It also facilitates interpretation alongside studies using different pathological definitions without implying that those definitions can be retrospectively harmonized. Reporting both measures prevents residual nodal disease from being obscured when primary-tumor response is used to summarize pathological activity.

The absence of survival events among patients achieving pCR and the favorable MPR and R0 associations were consistent with observational evidence linking pathological response with prognosis[22]. They did not validate pCR or MPR as survival surrogates[23]. Response and margin status were determined after treatment and incorporate tumor biology, treatment sensitivity, surgical factors and pathological assessment. Analyses measuring OS from enrollment might therefore be affected by surgical selection and the timing of ascertainment. TRG-adjusted CD3

models likewise conditioned on a post-treatment variable and did not establish causal independence. The dMMR subgroup comprised only 4 patients, its nominal pCR difference should be interpreted in the context of molecular heterogeneity and uncertain chemotherapy effects in dMMR disease[24].

Long-term follow-up provides additional descriptive information, although comparisons with external controls require careful alignment of eligibility criteria, time origins and endpoint definitions[25]. Early perioperative studies, including Neo-PLANET, evaluated different multimodality regimens and follow-up schedules[26]. Our data showed that some patients remained disease free for several years, while residual relapse risk remained apparent despite pathological regression. Late Kaplan-Meier estimates depend on progressively smaller risk sets and should be interpreted with their CIs and numbers at risk. Neither an unreached median nor a plateau at the end of follow-up establishes a cure fraction. Recurrence ascertainment and vital-status follow-up were distinct, particularly when reporting DFS and OS from different time origins. The sustained descriptive survival estimates are clinically informative, but prospective comparison remains necessary to determine whether the regimen improves outcomes over an appropriate contemporary control.

The observed safety profile was characterized mainly by hematologic and gastrointestinal events, consistent with the known toxicity profile of FLOT-based perioperative therapy[5]. The observed 72.5% grade ≥3 event rate was numerically close to the grade 3-4 rates reported in MATTERHORN (71.6% and 71.2%), but this does not establish equivalent safety[8]. Similarly, grade ≥3 TRAEs were reported in 65% and 63% of patients in the pembrolizumab and placebo groups of the final KEYNOTE-585 analysis[7]. NEOSUMMIT-01 used oxaliplatin-based doublets rather than FLOT and reported lower grade 3-4 treatment-related toxicity[17]. These comparisons provide context for the chemotherapy backbone but do not permit the regimens to be ranked by safety. Observation windows, cumulative exposure, event grading and causality attribution may differ. Incomplete postoperative treatment also means that toxicity should be interpreted together with actual treatment delivery rather than the planned regimen alone. Transparent harms reporting requires consistent definitions and complete capture of serious,

immune-related and treatment-limiting events[27]. The recorded endocrine, dermatological and pulmonary events identify areas requiring careful attribution; they do not show that every event was immune mediated.

Pretreatment immune profiling adds a biological perspective, and multi-omics studies indicate that perioperative treatment can remodel the tumor microenvironment[28]. None of the 36 individual features remained significantly different between TRG0-1 and TRG2-3 after correction for multiple testing. This finding is consistent with the complexity of immunotherapy response in GC[29]. PANDA identified higher baseline PD-1-positive CD8-positive T-cell infiltration in responders and reported dynamic immune remodeling after treatment[19]. Similarly, Chen and colleagues used multiplex immunohistochemistry and spatial information to show that a multidimensional tumor-infiltrating immune-cell signature was more informative for ICI response than a single immune marker. These data suggest that functional state, cellular interactions and spatial organization may be at least as important as the absolute abundance of an individual immune population[30]. In our TIME cohort, stromal CD3 positivity showed the most internally consistent survival signal. Patients with stromal CD3 positivity above the median had longer DFS and OS, and the association persisted after adjustment for TRG0-1 versus TRG2-3. This observation is consistent with prior evidence. A systematic review of 35 GC studies involving 4888 patients showed that high CD3-positive lymphocyte infiltration was associated with improved OS, particularly for CD3 infiltration in tumor stroma[31]. In stage II/III GC treated with fluoropyrimidine-based adjuvant chemotherapy, an immunoscore integrating CD3- and CD8-positive T-cell densities was independently associated with both DFS and OS[32]. A 2025 systematic review and meta-analysis likewise found that greater CD3-positive T-cell infiltration was associated with improved DFS in GC, whereas CD4- and CD8-positive populations showed compartment-dependent prognostic information[33]. Thus, higher stromal CD3 positivity may mark a pre-existing T-cell-inflamed immune contexture at the host-tumor interface that is associated with more effective antitumor immune control during subsequent multimodality treatment[34]. The recent study by Zhao et al. further illustrated the potential importance of immune-cell localization and functional heterogeneity[10]. In their EBV-associated

cohort, higher pretreatment CD8-positive T-cell infiltration was associated with MPR, predominantly within the tumor parenchyma rather than the stromal compartment. By contrast, our exploratory analysis did not identify any pretreatment immune feature that significantly distinguished MPR from non-MPR after correction for multiple testing. The most consistent nominal survival association involved stromal CD3 positivity. These observations are not necessarily contradictory because they concern different immune-cell populations, spatial compartments, clinical endpoints and molecularly distinct cohorts. CD3 positivity captures a broader T-cell population, whereas CD8-positive T cells more specifically represent a cytotoxic-enriched compartment. One hypothesis is that pre-existing intratumoral effector T-cell infiltration may be more closely related to immediate tumor regression, whereas stromal T-cell abundance may reflect a broader immune contexture associated with subsequent disease control and survival. This interpretation remains speculative. Zhao et al. also characterized treatment-associated immune changes using paired tissue specimens, including heterogeneous CD8-positive T-cell responses and reductions in stromal M1-like macrophage density. Our analysis was restricted to pretreatment specimens and therefore could not evaluate treatment-induced immune remodeling. Future studies incorporating longitudinal sampling, spatially resolved immune profiling, molecular stratification and independent validation are needed to determine whether these compartment-specific immune features have reproducible prognostic or treatment-predictive value.

A predictive biomarker requires evidence of differential treatment benefit by marker status, ideally from a randomized treatment-by-marker comparison[35]. This study had no FLOT-alone control arm. CD3 was selected and evaluated within the same 24-patient dataset, which included only 10 DFS events and 7 deaths. Repeating a common median-split procedure across the complete panel improves reporting transparency but does not create independent validation. Neither nominal log-rank significance nor a focused adjusted model overrides the nonsignificant FDR results. Median dichotomization also cannot substitute for continuous-effect assessment. Future studies should prioritize prespecified analyses, blinded image assessment, reproducible compartment segmentation and external cohorts, consistent with

lymphocyte-scoring and REMARK recommendations[36, 37]. Reproducibility across tissue blocks and laboratories is essential before the threshold can be considered for clinical use. Consistent marker definitions, positivity denominators and tissue sampling are particularly important when attempting to reproduce a percentage-based cut point from a small biopsy cohort.

Several limitations should be considered. First, this was a single-center, single-arm phase II trial with a modest sample size. Consequently, efficacy could not be separated from patient selection, surgical management, underlying tumor biology or secular changes in supportive care. Second, postoperative treatment exposure varied, and reasons for incomplete protocol therapy were incompletely documented for several patients, potentially affecting long-term outcomes. Third, although OS follow-up was mature, ascertainment of recurrence and progression was less uniformly updated than vital status, which might introduce uncertainty into DFS estimates. Fourth, the TIME cohort comprised only 24 of 51 patients and contained no known dMMR tumors, introducing tissue-availability and molecular-selection bias. Fifth, multiple immune features were investigated without an external validation cohort, the stromal CD3 analysis should therefore remain explicitly hypothesis-generating. Finally, archived pretreatment core biopsies could not fully capture the spatial heterogeneity of GC, and the absence of systematically paired post-treatment tissue limited assessment of treatment-induced immune remodeling.

In conclusion, perioperative toripalimab plus FLOT yielded durable long-term survival outcomes in this high-risk single-arm cohort with a ypT0N0 response of 13.7%, a TRG0 rate of 19.6% and a 5-year OS rate of 59.0%. Pretreatment immune profiling did not establish an FDR-confirmed pathological-response marker. Stromal CD3 positivity was the most consistent hypothesis-generating prognostic signal and should be validated prospectively before any clinical interpretation.

Methods

Study design and eligibility criteria

This prospective, open-label, single-arm phase II clinical trial was conducted at Tianjin

Medical University Cancer Institute and Hospital. Key protocol eligibility criteria were: (1) age 18-75 years; (2) previously untreated, histologically confirmed adenocarcinoma of the stomach or gastroesophageal junction; (3) locally advanced clinical stage II-III disease considered resectable with curative intent; (4) Eastern Cooperative Oncology Group (ECOG) performance status of 0-1; (5) adequate bone-marrow, hepatic, renal and cardiopulmonary function; and (6) measurable or otherwise evaluable disease according to Response Evaluation Criteria in Solid Tumours version 1.1 (RECIST 1.1). Main exclusion criteria included: (1) active autoimmune disease requiring systemic treatment, uncontrolled infection or previous organ transplantation; (2) clinically significant cardiopulmonary disease; and (3) another active malignancy or any condition that might make study treatment or major surgery unsafe. Further eligibility criteria are detailed in the Supplementary materials and study protocol.

The study protocol was approved by the institutional ethics committee of Tianjin Medical University Cancer Institute and Hospital (approval: E2019401). The study was conducted in accordance with the Declaration of Helsinki, Good Clinical Practice and applicable institutional requirements. All patients provided written informed consent for trial participation. Use of archived diagnostic tissue for exploratory translational analyses was covered by the approved research programme. The trial was registered at ClinicalTrials.gov (NCT04354662).

Treatment

Patients were scheduled to receive four 14-day cycles of neoadjuvant toripalimab plus FLOT. Toripalimab 240 mg was administered intravenously on day 1. FLOT comprised docetaxel 50 mg/m^2, oxaliplatin 85 mg/m^2 and leucovorin 200 mg/m^2 administered intravenously on day 1, followed by fluorouracil 2600 mg/m^2 as a 24-hour continuous infusion. Restaging and multidisciplinary reassessment preceded gastrectomy. Surgery aimed for complete resection and D2 lymphadenectomy when anatomically and clinically appropriate. Postoperative treatment, intended to begin after recovery and within approximately 2 months of surgery, consisted of four further 14-day cycles of the same combination. Dose delay, reduction, component discontinuation or cessation was permitted for toxicity or clinical need according to the protocol and investigator judgment.

Assessments and endpoints

Baseline evaluation included medical history, physical examination, performance status, laboratory testing, endoscopy with biopsy and cross-sectional imaging. Radiological response was classified according to RECIST 1.1. Objective response rate (ORR) and disease-control rate (DCR) were calculated only in the RECIST-evaluable subset. Surgery, resection type, lymphadenectomy category, margin status, postoperative stay, examined and positive lymph-node counts, and postoperative complications were abstracted from the clinical database and source-report summaries.

The co-primary endpoints were pCR and the 3-year DFS rate. Pathological response was assessed in the resected primary tumor according to the tumor regression grade (TRG) system, pCR was defined as TRG0, indicating complete regression with no viable tumor cells identified in the primary tumor bed. A more stringent nodal-inclusive pathological endpoint, defined as ypT0N0, was additionally reported and analyzed separately. Major pathological response (MPR) was defined as TRG0-1, representing complete or near-complete pathological regression. DFS was defined from the date of surgery to the first documented disease recurrence or progression, or death from any cause, whichever occurred first.

Secondary endpoints included MPR, R0 resection rate, D2 lymphadenectomy rate, OS and safety. OS was defined as the interval from study enrollment to death from any cause, patients who were alive at the data cutoff were censored at the date they were last known to be alive. Safety was characterized by the incidence and severity of treatment-related adverse events (TRAEs) occurring during perioperative treatment. TRAEs were graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE), version 5.0. Adverse events were reported as any-grade and grade ≥3 events. Exploratory translational endpoints evaluated associations between pretreatment tumor immune microenvironment (TIME) features, pathological regression, and long-term survival outcomes in the 24 patients with evaluable pretreatment tumor biopsies.

Translational TIME analysis

Quality-controlled pretreatment formalin-fixed, paraffin-embedded (FFPE) core-biopsy

specimens from 24 patients were used for multiplex immunofluorescence-based tumor immune microenvironment (TIME) profiling. Multiplex staining was performed using the PANO 7-plex immunohistochemistry kit (catalogue no. 0004100100; Panovue, Beijing, China) according to the manufacturer's protocol and a sequential tyramide signal amplification workflow. Briefly, tissue sections were deparaffinized, rehydrated and subjected to heat-induced antigen retrieval using citrate-based retrieval buffer. After endogenous peroxidase blocking, sequential rounds of primary-antibody incubation, horseradish peroxidase-conjugated secondary-antibody detection and tyramide signal amplification were performed. Heat-mediated antigen retrieval and antibody stripping were carried out between staining cycles. After multiplex staining, nuclei were counterstained with 4′,6-diamidino-2-phenylindole (DAPI).

The principal immune panel included PD-1, PD-L1, CD3, CD8 and pan-cytokeratin (panCK). Additional multiplex measurements generated CD56-positive natural killer cell and CD68/CD163-defined macrophage phenotypes. Slides were scanned using the Mantra Quantitative Pathology Imaging System (PerkinElmer), and multispectral images were analyzed using inForm image-analysis software (PerkinElmer). PanCK staining distinguished tumor-parenchymal from stromal compartments. Within each compartment, immune-cell positivity was defined as the proportion of marker-positive cells among the corresponding analyzed cell population, whereas cell density was expressed as the number of marker-positive cells per unit tissue area (cells/mm^2). CD68-positive/CD163-negative and CD68-positive/CD163-positive macrophages were operationally categorized as M1-like and M2-like phenotypes, respectively. These labels were phenotypic descriptors and were not intended to imply functional macrophage polarization.

Patient-level TIME features included the densities and/or positive-cell proportions of CD3, CD8, PD-1, PD-L1, CD56, CD3-positive/PD-1-positive, CD8-positive/PD-1-positive and CD68/CD163-defined macrophage populations within tumor-parenchymal and stromal compartments. For heatmap visualization, immune variables were log1p-transformed and standardized within each feature to generate Z-scores. Quantitative analyses used digitally derived TIME measurements; representative multiplex images were used only for illustration.

Immune-marker survival associations were assessed using log1p-transformed, standardized univariable Cox models. Time-dependent AUCs at 36 and 60 months were estimated using a censored-survival ROC framework. Kaplan–Meier candidate plots used the cohort median rather than an optimized cut point. Interpretation followed REMARK principles.[36] The 24-patient TIME subset was compared descriptively with the remaining 27 patients to assess tissue-availability selection. All biomarker analyses were exploratory. The median split for stromal CD3 positivity was used for visualization and was not proposed as a clinically validated threshold.

Sample size calculation

Sample size was estimated using Simon's optimal two-stage design with pCR as the primary endpoint. A one-sided type I error rate (α) of 0.05 and a type II error rate (β) of 0.20 were specified, corresponding to 80% power. Previous studies of neoadjuvant or perioperative FLOT in patients with resectable GC reported heterogeneous pCR rates ranging from approximately 7% to 20%.[38] Accordingly, the null pCR rate (P0) was set at 7%, representing the historical reference rate, and a pCR rate of 20% (P1) was considered clinically promising for toripalimab combined with FLOT. The sample size was calculated using the Two-Stage Phase II Clinical Trials procedure in PASS version 15.0. Under Simon's optimal two-stage design, 16 patients were required in the first stage. If one or fewer patients achieved pCR among the first 16 patients, the study would terminate early for futility. If at least two patients achieved pCR, the study would proceed to the second stage, in which an additional 34 patients would be enrolled. At the end of both stages, if at least 7 of 50 patients achieved pCR, the null hypothesis would be rejected and the regimen would be considered sufficiently promising to warrant further investigation. The planned sample size was therefore 50 patients.

Statistical analysis

Continuous variables were summarized as mean and standard deviation (SD) or median, interquartile range (IQR) and range, as appropriate. Categorical variables were summarized as counts and percentages. Exact two-sided 95% confidence intervals (CIs) for binomial proportions, including pCR, MPR, ORR and DCR were calculated using the Clopper-Pearson

method. For protocol-defined binary pathological endpoints, all 51 treated patients were retained in the denominator; ORR and DCR were calculated in the 9 patients with measurable disease at baseline. OS and DFS were estimated using Kaplan-Meier methods with pointwise 95% CIs and numbers at risk in Stata version 19.0 (StataCorp LLC, College Station, TX, USA). The patient-level swimmer plot depicting neoadjuvant and postoperative treatment exposure, surgery, disease recurrence or progression, death and follow-up was generated using PROC SGPLOT in SAS version 9.4 (SAS Institute Inc., Cary, NC, USA). Median follow-up was estimated using the reverse Kaplan-Meier method. Because this was a single-arm study, efficacy estimates are descriptive, and no causal comparison with historical or contemporary trials was attempted.

For the TIME analysis, marker distributions were compared between the recorded TRG0-1 and TRG2-3 groups using non-parametric tests. Receiver-operating-characteristic (ROC) analysis described discrimination for pathological regression. Associations between individual TIME features and DFS or OS were explored using univariable Cox proportional-hazards models after log1p transformation and standardization per standard deviation. As a post hoc sensitivity analysis, each of the 36 raw TIME features was dichotomized at its respective cohort median (>median versus ≤median) and evaluated using univariable Cox regression and two-sided log-rank tests. Benjamini-Hochberg false-discovery-rate (FDR) correction was applied separately across the 36 comparisons for DFS and OS. Stromal CD3 positivity was visualized in the main Kaplan-Meier analyses using its cohort median of 25.985%. All median-split analyses were exploratory and were not intended to establish clinically validated thresholds. Time-dependent discrimination at 3 and 5 years was estimated using inverse-probability-of-censoring weighting. Exploratory subgroup analyses of protocol-defined pCR were performed using Fisher's exact test. Subgroup analyses of DFS and OS used the Kaplan-Meier method, log-rank test and univariable Cox proportional-hazards models. Given the limited number of events, multivariable analyses were intentionally parsimonious. In the full cohort, MPR and R0 resection were entered simultaneously into multivariable Cox models for DFS and OS. In the TIME cohort, an exploratory two-variable Cox model included the stromal CD3 positivity

group and pathological response category (TRG0-1 versus TRG2-3). Because of the small number of events in the TIME cohort, profile-likelihood CIs and likelihood-ratio *P* values were used. All subgroup and multivariable analyses were considered exploratory. $P<0.05$ was considered statistically significant.

Declarations

Ethics approval and consent to participate

The study was approved by the institutional ethics committee of Tianjin Medical University Cancer Institute and Hospital (E2019401). All patients were provided written informed consent for trial participation.

Trial registration

This study was registered at ClinicalTrials.gov (NCT04354662). The registration record was first registered on 15 April 2020 and first posted on 21 April 2020.

Funding

This study was supported by the Tianjin Key Medical Discipline Construction Project (Grant No. TJYXZDXK-3-003A).

Conflicts of interest

The authors declare that there are no conflicts of interest.

Author contributions

Hong-Li Li: Conceptualization, methodology, investigation, project administration and writing-original draft. Shao-Hua Ge: Methodology, formal analysis, data curation, visualization and writing-original draft. Feng-Lin Zang: Methodology, investigation, validation, visualization and writing-original draft.

Le Zhang, Yu-Chong Yang, Jing-Jing Duan, Peng Ren, Fu-Liang Cao, Ming Bai, Tao Ning, Li-Kun Zhou, Fei-Xue Wang, Jia-Yu Zhang, Yan-Sha Sun, Rui Liu, Yi Ba, Li-Yan Zhou, Xiao-Qing Li, Yue Song, Yi-Ran Si and Yi-Lin Deng: Investigation, resources, data curation and writing-review and editing.

Jing-Yu Deng: Conceptualization, methodology, supervision and writing-review and editing.

Ting Deng: Conceptualization, supervision, project administration and writing-review and editing.

All authors reviewed and approved the final manuscript and agreed to be accountable for the work.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgments

We thank all patients and their relatives and all members of the collaborative group in this trial. We also thank Daoyu Zhang (Jiangsu Hengrui Pharmaceuticals) for data interpretation and medical writing support.

628 **Table 1. Baseline characteristics of 51 patients with G/GEJ adenocarcinoma**

Baseline characteristics	Total (N=51)	Percentage
Age (years)		
Median (range)	59 (28-75)	
>59	23	45.1%
≤59	28	54.9%
Sex		
Male	38	74.5%
Female	13	25.5%
ECOG performance status		
0	25	49.0%
1	26	51.0%
Primary tumor site		
Gastric	45	88.2%
Gastroesophageal junction	2	3.9%
Other	4	7.8%
Clinical T stage		
T2	2	3.9%
T4	48	94.1%
Unknown	1	2.0%
Clinical N stage		
N0	5	9.8%
N1	28	54.9%
N2	13	25.5%
N3	3	5.9%
Unknown	2	3.9%
Clinical stage at baseline (AJCC 8th edition)		
IIA	1	2.0%
IIB	6	11.8%
IIIA	27	52.9%
IIIB	11	21.6%
IIIC	3	5.9%
Unknown	3	5.9%
HER2 expression		
0	31	60.8%
1+	5	9.8%
2+	11	21.6%
Unknown	4	7.8%
Mismatch repair status		
pMMR	45	88.2%
dMMR	4	7.8%
Unknown	2	3.9%
EBER status		
Negative	42	82.4%
Positive	6	11.8%
Unknown	3	5.9%

629 Data are n (%) unless otherwise indicated.
630 AJCC, American Joint Committee on Cancer; dMMR, deficient mismatch repair; EBER, Epstein-Barr virus-encoded
631 RNA; ECOG, Eastern Cooperative Oncology Group; GEJ, gastroesophageal junction; HER2, human epidermal
632 growth factor receptor 2; pMMR, proficient mismatch repair.
633

Table 2. Safety profile of perioperative toripalimab plus FLOT

Treatment-related adverse events	Any grade, n (%)	Grade ≥3, n (%)
Any TRAE	**51 (100.0)**	**37 (72.5)**
Neutropenia	46 (90.2)	26 (51.0)
Leukopenia	44 (86.3)	14 (27.5)
Nausea	38 (74.5)	0
Anemia	37 (72.5)	3 (5.9)
Fatigue	32 (62.7)	0
Alopecia	32 (62.7)	0
Decreased appetite	28 (54.9)	0
Dizziness	22 (43.1)	0
Increased ALT/AST	19 (37.3)	3 (5.9)
Hypoalbuminemia	19 (37.3)	1 (2.0)
Peripheral sensory neuropathy	17 (33.3)	0
Vomiting	15 (29.4)	0
Thrombocytopenia	15 (29.4)	0
Diarrhea	12 (23.5)	0
Hyperglycemia	10 (19.6)	0
Headache	9 (17.6)	0
Rash	8 (15.7)	3 (5.9)
Oral mucositis	7 (13.7)	2 (3.9)
Pyrexia	7 (13.7)	0
Lymphopenia	7 (13.7)	0
Hypoglycemia	6 (11.8)	0
Hypokalemia	5 (9.8)	2 (3.9)
Increased bilirubin	3 (5.9)	1 (2.0)
Hypothyroidism	3 (5.9)	0
Bone marrow suppression	1 (2.0)	1 (2.0)
Hyperthyroidism	1 (2.0)	0
Pneumonitis	1 (2.0)	0

Data are presented as n (%). The safety analysis set included all 51 patients who received at least one dose of study treatment. For each TRAE, the highest grade experienced by each patient during perioperative treatment was used. Exact two-sided 95% confidence intervals for the overall incidence of any-grade and grade ≥3 TRAEs were calculated using the Clopper-Pearson method.

ALT, alanine aminotransferase; AST, aspartate aminotransferase; CI, confidence interval; FLOT, fluorouracil, leucovorin, oxaliplatin, and docetaxel; NE, not evaluable; TRAE, treatment-related adverse event.

Figure legends

Figure 1. Study flow and treatment disposition.

Of 55 screened patients, 4 were excluded and 51 received toripalimab plus FLOT and underwent gastrectomy. At data cutoff, 36 patients had completed planned perioperative treatment and 15 had not. All 51 patients were included in the full and safety analysis sets, and 24 in the pretreatment tumor immune microenvironment (TIME) analysis set.

Abbreviations: FLOT, fluorouracil, leucovorin, oxaliplatin, and docetaxel; TIME, tumor immune microenvironment.

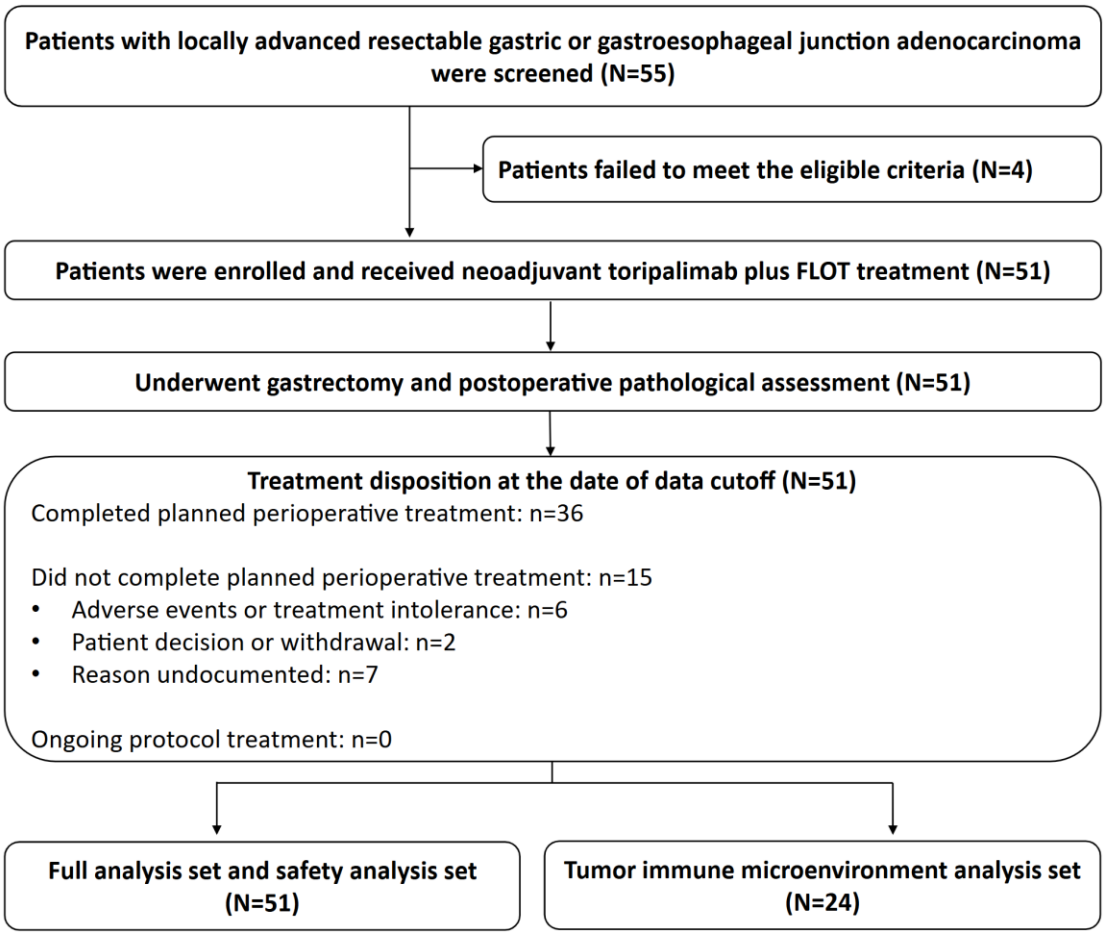


Figure 2. Pathological response and perioperative treatment exposure.

(A) Protocol-defined pathological complete response (pCR; TRG0) and major pathological response (MPR; TRG0-1) rates among 51 patients. Error bars indicate exact two-sided 95% confidence intervals calculated using the Clopper-Pearson method. (B) Swimmer plot showing

perioperative treatment, surgery and follow-up for each patient. Blue lines indicate alive/censored follow-up, and red lines indicate follow-up ending in death. Teal and orange segments represent neoadjuvant and postoperative adjuvant treatment, respectively. Diamonds indicate surgery, orange triangles indicate first recurrence or progression, red crosses indicate death, vertical ticks indicate censoring, purple stars indicate inclusion in the tumor immune microenvironment (TIME) analysis (n=24), and the open square indicates a missing postoperative treatment end date (n=1).

Abbreviations: MPR, major pathological response; pCR, pathological complete response; TIME, tumor immune microenvironment; TRG, tumor regression grade.

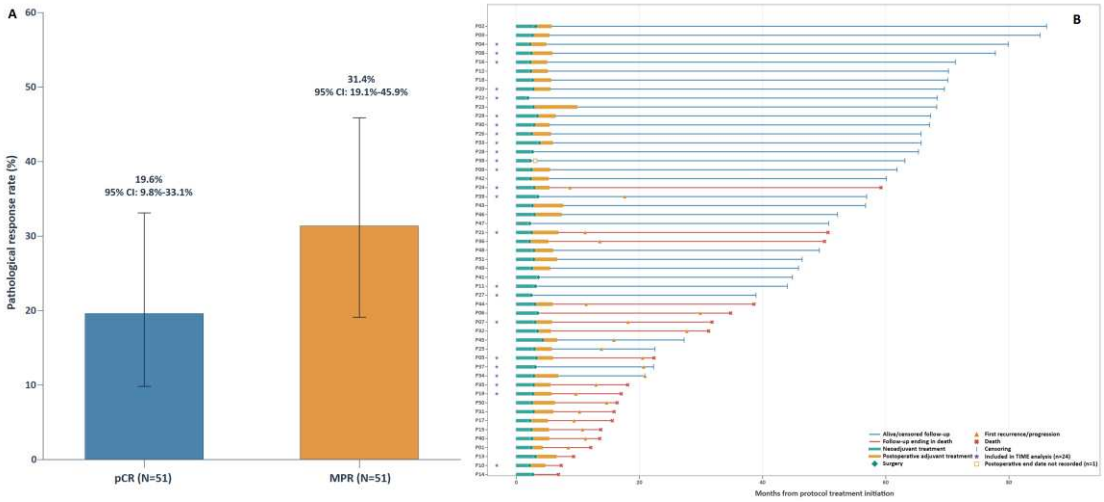


Figure 3. Long-term overall survival and disease-free survival outcomes.

(A) Kaplan-Meier curve for OS in the full analysis set (N = 51). (B) Kaplan-Meier curve for DFS in the same cohort. Vertical tick marks indicate censored observations. Numbers at risk and cumulative numbers censored are shown below each panel. Median survival and 24-, 36- and 60-month survival estimates with 95% CIs are displayed within the plots.

Abbreviations: CI, confidence interval; DFS, disease-free survival; NR, not reached; OS, overall survival.

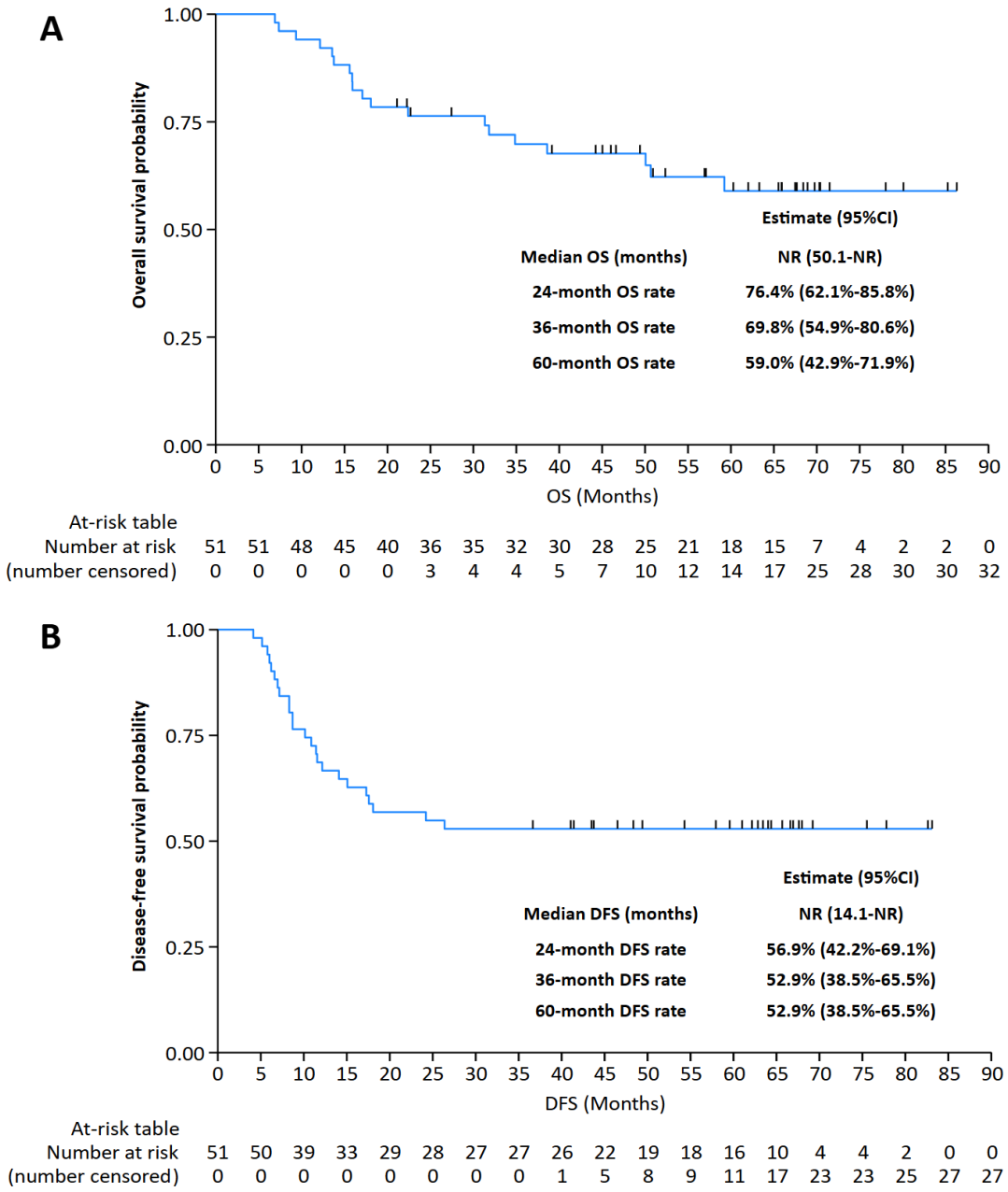


Figure 4. Pretreatment tumor immune microenvironment and prognostic relevance of stromal CD3 positivity.

(A) Heatmap of 36 pretreatment TIME features in 24 patients. Patients were ordered according to tumor regression grade (TRG0-TRG3). Immune features were log1p-transformed and standardized within each feature to generate row Z-scores. Colors therefore represent relative abundance across patients for each individual immune feature rather than absolute differences between different markers. (B) Distribution of stromal CD3 positivity according to pathological regression group. Violin plots show the distribution of individual patient values with embedded

box plots indicating the median and interquartile range. *P* value was calculated using a two-sided Mann-Whitney U test. (C) Kaplan-Meier curves for OS according to stromal CD3 positivity. (D) Kaplan-Meier curves for DFS according to stromal CD3 positivity. For survival analyses, patients were dichotomized at the cohort median stromal CD3 positivity of 25.985% into CD3-high (>25.985%; n=12) and CD3-low (≤25.985%; n=12) groups. Vertical tick marks indicate censored observations, and numbers at risk are shown below the curves. *P* values in panels C and D were calculated using the log-rank test.

Abbreviations: DFS, disease-free survival; OS, overall survival; PD-1, programmed cell death protein 1; PD-L1, programmed death-ligand 1; TIME, tumor immune microenvironment; TRG, tumor regression grade.

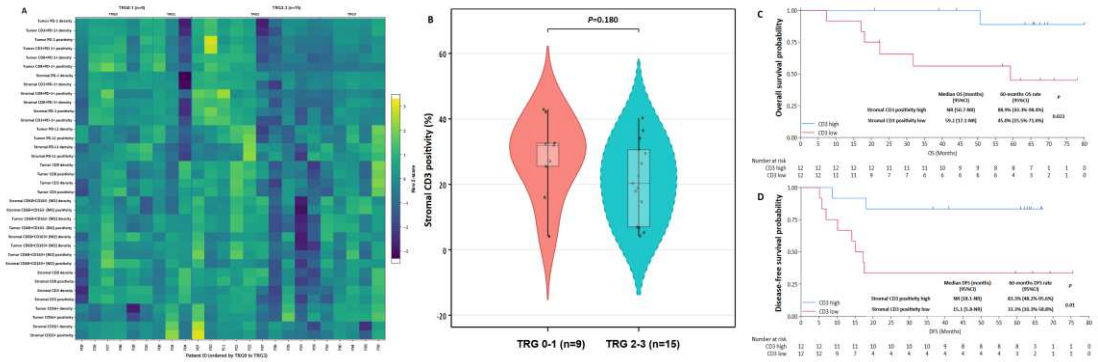
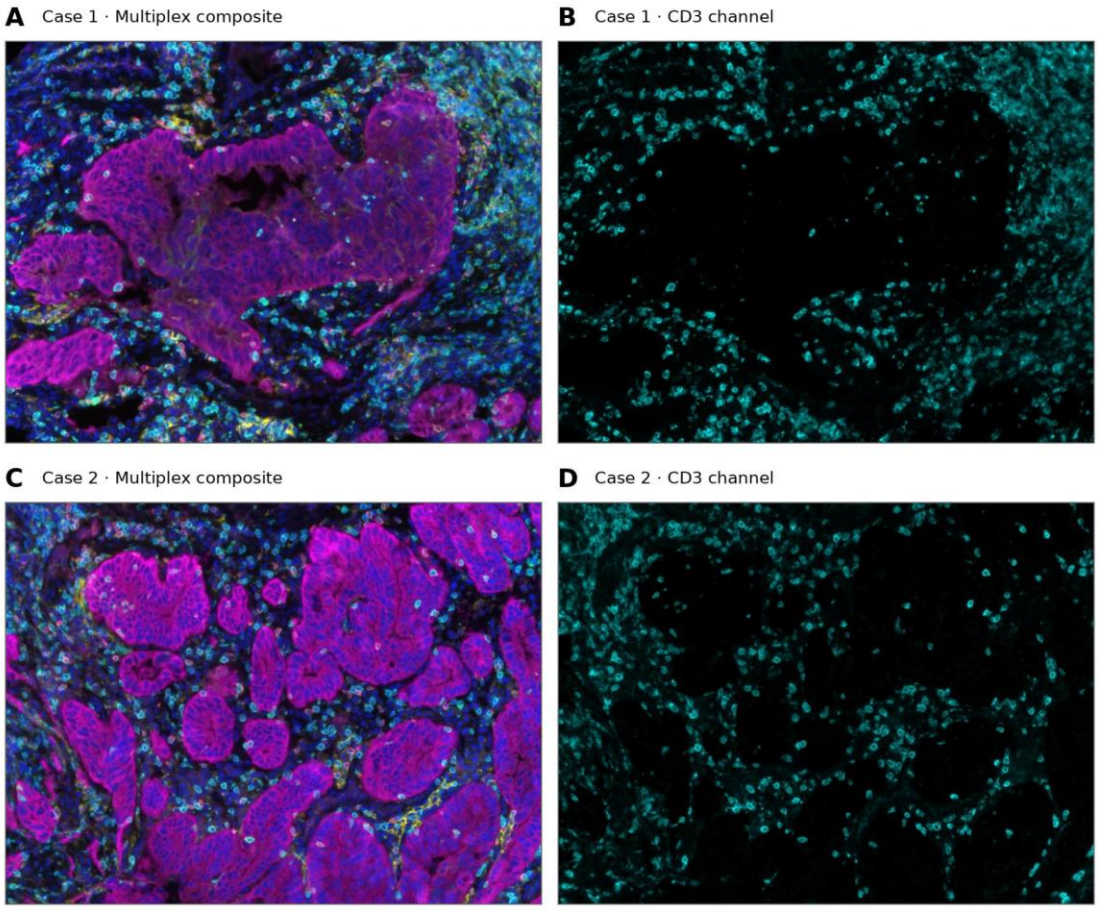


Figure 5. Representative multiplex immunofluorescence images illustrating stromal CD3-positive T-cell infiltration in pretreatment gastric tumor biopsies.

Representative pretreatment multiplex immunofluorescence images from two patients classified in the stromal CD3-high subgroup. (A) Multiplex composite image from representative case 1. (B) Corresponding CD3 single-channel image from case 1, illustrating abundant CD3-positive lymphocytes within the stromal compartment surrounding tumor structures. (C) Multiplex composite image from representative case 2. (D) Corresponding CD3 single-channel image from case 2. Multiplex staining was performed on formalin-fixed, paraffin-embedded pretreatment core-biopsy specimens and pan-cytokeratin was included in the multiplex panel to support distinction between tumor-parenchymal and stromal compartments. Images illustrate the assay phenotype and should not be interpreted as evidence

of between-group differences. Quantitative stromal CD3 positivity used in the survival analyses was derived from digital image analysis rather than visual scoring of these representative fields. Abbreviations: CD3, cluster of differentiation 3; FFPE, formalin-fixed paraffin-embedded; TIME, tumor immune microenvironment.



Reference

1. Bray F*, et al.* Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* **74**, 229-263 (2024).

2. Zheng R*, et al.* Cancer incidence and mortality in China, 2024. *Journal of the National Cancer Center*, (2026).

3. Shitara K*, et al.* Pan-Asian adapted ESMO Clinical Practice Guidelines for the diagnosis, treatment and follow-up of patients with gastric cancer. *ESMO Open* **9**, 102226 (2024).

4. Li C*, et al.* Pathologic Response of Phase III Study: Perioperative Camrelizumab Plus Rivoceranib and Chemotherapy Versus Chemotherapy for Locally Advanced Gastric Cancer (DRAGON IV/CAP 05). *J Clin Oncol* **43**, 464-474 (2025).

5. Al-Batran SE*, et al.* Perioperative chemotherapy with fluorouracil plus leucovorin, oxaliplatin, and docetaxel versus fluorouracil or capecitabine plus cisplatin and epirubicin for locally advanced, resectable gastric or gastro-oesophageal junction adenocarcinoma (FLOT4): a randomised, phase 2/3 trial. *Lancet* **393**, 1948-1957 (2019).

6. Lorenzen S*, et al.* Perioperative Atezolizumab Plus Fluorouracil, Leucovorin, Oxaliplatin, and Docetaxel for Resectable Esophagogastric Cancer: Interim Results From the Randomized, Multicenter, Phase II/III DANTE/IKF-s633 Trial. *J Clin Oncol* **42**, 410-420 (2024).

7. Shitara K*, et al.* Pembrolizumab Plus Chemotherapy Versus Chemotherapy as Perioperative Therapy in Locally Advanced Gastric and Gastroesophageal Junction Cancer: Final Analysis of the Randomized, Phase III KEYNOTE-585 Study. *J Clin Oncol* **43**, 3152-3159 (2025).

8. Janjigian YY*, et al.* Perioperative Durvalumab in Gastric and Gastroesophageal Junction Cancer. *N Engl J Med* **393**, 217-230 (2025).

9. Nie RC*, et al.* Perioperative Toripalimab Plus Chemotherapy Versus Chemotherapy Alone in Locally Advanced Gastric or Gastroesophageal Junction Cancer: 3-Year Follow-Up of NEOSUMMIT-01 Trial. *J Clin Oncol* **44**, 1869-1875 (2026).

10. Zhao L*, et al.* Neoadjuvant toripalimab plus CapeOX in locally advanced Epstein-Barr virus-associated gastric or gastroesophageal junction adenocarcinoma: a phase Ⅱ trial. *Nat Commun* **17**, (2026).

11. Omori R, Fujiwara Y, Tokunaga K, Sato T, Mukherjee S. Perioperative chemoimmunotherapy for patients with gastric or gastroesophageal junction cancer: a systematic review and meta-analysis. *JNCI Cancer Spectr* **10**, (2026).

12. Möhring C*, et al.* Perioperative therapy with FLOT4 significantly increases survival in

patients with gastroesophageal and gastric cancer in a large real-world cohort. *Int J Cancer* **153**, 609-622 (2023).

13. Zhao Y*, et al.* Comprehensive multi-omics analysis of resectable locally advanced gastric cancer: Assessing response to neoadjuvant camrelizumab and chemotherapy in a single-center, open-label, single-arm phase II trial. *Clin Transl Med* **14**, e1674 (2024).

14. Shitara K*, et al.* Neoadjuvant and adjuvant pembrolizumab plus chemotherapy in locally advanced gastric or gastro-oesophageal cancer (KEYNOTE-585): an interim analysis of the multicentre, double-blind, randomised phase 3 study. *Lancet Oncol* **25**, 212-224 (2024).

15. Janjigian YY*, et al.* Perioperative durvalumab plus fluorouracil, leucovorin, oxaliplatin, and docetaxel for resectable gastric and gastro-oesophageal junction adenocarcinoma (MATTERHORN): final results of overall survival and event-free survival by pathological response in a global, randomised, double-blind, placebo-controlled, multicentre, phase 3 trial. *Lancet* **408**, 1114-1128 (2026).

16. Ilson DH. KEYNOTE-585 Fails While Matterhorn Succeeds in Gastric Cancer: What Lessons Can We Learn? *J Clin Oncol* **43**, 3141-3143 (2025).

17. Yuan SQ*, et al.* Perioperative toripalimab and chemotherapy in locally advanced gastric or gastro-esophageal junction cancer: a randomized phase 2 trial. *Nat Med* **30**, 552-559 (2024).

18. Rodrigues LLC*, et al.* Perioperative Chemoimmunotherapy vs. Chemotherapy Alone for Gastroesophageal Cancer: A Systematic Review and Meta-Analysis. *J Gastrointest Cancer* **57**, 40 (2026).

19. Verschoor YL*, et al.* Neoadjuvant atezolizumab plus chemotherapy in gastric and gastroesophageal junction adenocarcinoma: the phase 2 PANDA trial. *Nat Med* **30**, 519-530 (2024).

20. Tong Y, Zhu Y, Zhao Y, Shan Z, Liu D, Zhang J. Evaluation and Comparison of Predictive Value of Tumor Regression Grades according to Mandard and Becker in Locally Advanced Gastric Adenocarcinoma. *Cancer Res Treat* **53**, 112-122 (2021).

21. Zhu K, Jin H, Zhang Q, Shou C, Chen F, Yu J. ypT0 gastric carcinoma after preoperative chemotherapy: a unique status according to AJCC 8(th) edition cancer staging system. *Transl Cancer Res* **9**, 7384-7393 (2020).

22. Airoldi M*, et al.* Prognostic impact of pathological complete response and response-adapted outcomes in patients with gastroesophageal adenocarcinoma treated with neoadjuvant or perioperative treatments: a multicentric cohort study. *ESMO Open* **11**, 106084 (2026).

23. Leite LF*, et al.* Is pathological complete response a surrogate for overall survival in neoadjuvant clinical trials for gastroesophageal adenocarcinoma? *Eur J Cancer* **237**, 116602 (2026).

24. Liu B*, et al.* Perioperative chemotherapy for gastric cancer patients with microsatellite instability or deficient mismatch repair: A systematic review and meta-analysis. *Cancer* **131**, e35831 (2025).

25. Seeger JD*, et al.* Methods for external control groups for single arm trials or long-term uncontrolled extensions to randomized clinical trials. *Pharmacoepidemiol Drug Saf* **29**, 1382-1392 (2020).

26. Tang Z*, et al.* The Neo-PLANET phase II trial of neoadjuvant camrelizumab plus concurrent chemoradiotherapy in locally advanced adenocarcinoma of stomach or gastroesophageal junction. *Nat Commun* **13**, 6807 (2022).

27. Junqueira DR*, et al.* CONSORT Harms 2022 statement, explanation, and elaboration: updated guideline for the reporting of harms in randomised trials. *Bmj* **381**, e073725 (2023).

28. Ji Z*, et al.* Multiomics reveals tumor microenvironment remodeling in locally advanced gastric and gastroesophageal junction cancer following neoadjuvant immunotherapy and chemotherapy. *J Immunother Cancer* **12**, (2024).

29. Chen Y*, et al.* Predicting response to immunotherapy in gastric cancer via multi-dimensional analyses of the tumour immune microenvironment. *Nat Commun* **13**, 4851 (2022).

30. Sun K*, et al.* scRNA-seq of gastric tumor shows complex intercellular interaction with an alternative T cell exhaustion trajectory. *Nat Commun* **13**, 4943 (2022).

31. Jiang W*, et al.* Tumor-infiltrating immune cells and prognosis in gastric cancer: a systematic review and meta-analysis. *Oncotarget* **8**, 62312-62329 (2017).

32. Yun S*, et al.* Immunoscore is a strong predictor of survival in the prognosis of stage II/III gastric cancer patients following 5-FU-based adjuvant chemotherapy. *Cancer Immunol Immunother* **70**, 431-441 (2021).

33. Cao X*, et al.* Prognostic role of tumor-infiltrating lymphocytes in gastric cancer: A systematic review and meta-analysis. *Clin Res Hepatol Gastroenterol* **49**, 102510 (2025).

34. Jiang Y*, et al.* ImmunoScore Signature: A Prognostic and Predictive Tool in Gastric Cancer. *Ann Surg* **267**, 504-513 (2018).

35. Ballman KV. Biomarker: Predictive or Prognostic? *J Clin Oncol* **33**, 3968-3971 (2015).

36. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. REporting recommendations for tumour MARKer prognostic studies (REMARK). *Br J Cancer* **93**, 387-391 (2005).

37. Hendry S*, et al.* Assessing Tumor-infiltrating Lymphocytes in Solid Tumors: A Practical Review for Pathologists and Proposal for a Standardized Method From the International Immunooncology Biomarkers Working Group: Part 1: Assessing the Host

Immune Response, TILs in Invasive Breast Carcinoma and Ductal Carcinoma In Situ, Metastatic Tumor Deposits and Areas for Further Research. *Adv Anat Pathol* **24**, 235-251 (2017).

38. Al-Batran SE*, 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* **17**, 1697-1708 (2016).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementarymeterialsTablesandFigures.docx
- SupplementarymeterialsStudyprotocol.docx
- SupplementarymetarialsStatisticalAnalysisPlan.docx