ABSTRACT

Objectives Some studies have identified tumour-infiltrating lymphocytes (TILs) in H&E-stained sections of gastric cancer, but the prognostic and clinicopathological significance of this remains unclear. The objective of this study is to evaluate the associations between H&E-based TIL density and prognosis and clinicopathological characteristics of patients with gastric cancer.

Design Systematic review and meta-analysis.

Data sources Cochrane Library, PubMed and Embase databases were searched through 25 February 2020.

Eligibility criteria Studies evaluating the correlations between TILs assessed by H&E-stained sections and prognosis and clinicopathological characteristics of gastric cancer were included.

Data extraction and synthesis Relevant data were extracted and risks of bias were assessed independently by two reviewers. HR and relative risk (RR) with 95% CI were pooled by random-effect models to estimate the associations between TIL density and overall survival (OS) and clinicopathological characteristics, respectively.

Results We enrolled nine studies including 2835 cases for the present meta-analysis. High TILs were associated with superior OS (HR=0.68, 95% CI 0.52 to 0.87, p=0.003) compared with low TILs. High TILs were significantly associated with lower depth of invasion (T3–T4 vs T1–T2) (RR=0.58, 95% CI 0.50 to 0.66, p<0.001), less lymph node involvement (presence vs absence) (RR=0.68, 95% CI 0.56 to 0.81, p<0.001) and earlier TNM (tumour, node, metastasis) stage (III–IV vs I–II) (RR=0.68, 95% CI 0.55 to 0.83, p<0.001). TIL density was not associated with age, gender, Lauren classification or histological grade. The methodology for evaluating TIL and its cut-off value varied across different studies, which might affect the results of our meta-analysis.

Conclusions Our meta-analysis suggests that H&E-based TIL density is a reliable biomarker to predict the clinical outcomes of patients with gastric cancer. Multicentre, prospective studies are needed to further confirm our findings.

PROSPERO registration number CRD42020169877.

INTRODUCTION

Gastric cancer ranks fifth among the most frequently diagnosed carcinoma and third among the most common causes of cancer-related deaths worldwide.1 Nearly 60% of patients with gastric cancer present with late-stage cancer at diagnosis.2 Over 40% of patients who underwent curative resection experienced recurrence within 2 years after operation.3 Despite improvement in the diagnosis and therapy, the prognosis of patients with gastric cancer remains dismal, with 5-year overall survival (OS) rates ranging from 28% to 51%.4–7

Among multiple clinicopathological features, such as tumour size, histological grade, Lauren classification, lymphovascular invasion, and overexpression or amplification of the human epidermal growth factor receptor 2, tumour, node, metastasis (TNM) staging system is regarded as a key predictor of clinical outcome in patients with gastric cancer.8 However, heterogeneity exists in the survival outcomes of patients with the same TNM stage.9 There is an unmet need to identify new biomarkers to refine the prognosis of patients. Accumulating evidence suggests tumour microenvironment (TME) (composed of fibroblasts, extracellular matrix, endothelial cells and immune cells) as a key contributor in tumourigenesis, growth, invasion and metastasis.10 The assessment

Strengths and limitations of this study

To our knowledge, this is the first meta-analysis to evaluate the prognostic role of tumour-infiltrating lymphocyte (TIL) density evaluated by H&E-stained sections in gastric cancer.

This meta-analysis was performed by following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA).

H&E-based TIL assessment cannot provide information about subsets of TILs.

Heterogeneity among different studies is inevitable.
of tumour microenvironmental features seems to have considerable prognostic value.\textsuperscript{11}

Tumour-infiltrating lymphocytes (TILs), including lymphocytes and plasma cells, reflect the host immune response to cancers and are associated with the prognosis of various solid tumours.\textsuperscript{12–19} The Immunoscore,\textsuperscript{14 15 20 21} which is a scoring system according to the density of CD3\textsuperscript{+} and CD8\textsuperscript{+} T cells at the tumour centre and the invasive margin, has been demonstrated globally in stage I–III colon cancer as a superior prognostic marker than TNM staging system, lymphovascular invasion, grade of differentiation and microsatellite instability status.\textsuperscript{15} High TIL levels significantly increased the sensitivity and efficacy of neoadjuvant chemotherapy in patients with breast cancer,\textsuperscript{16 17} and were positively associated with survival outcomes in patients with primary cutaneous melanoma, non-small cell lung cancer or breast cancer.\textsuperscript{17–19}

To date, there are only a few meta-analyses assessing the relationship between TILs and gastric cancer.\textsuperscript{22–24} Most studies included in these meta-analyses evaluated TIL subsets by immunohistochemistry (IHC). However, IHC-based TIL evaluation cannot provide comprehensive information due to the heterogeneity of TILs in different locations, and the use of IHC requires additional processing and technical standardisation. Compared with IHC-based TIL evaluation, H&E-based TIL assessment is inexpensive and applicable to routine pathology practice and clinical trials, even in resource-poor settings. The International Immuno-Oncology Biomarkers Working Group (IIOBWG) recommends a standardised method for TIL evaluation based on an H&E-stained section in breast cancer.\textsuperscript{25} To date, there is no consensus on H&E-based TIL evaluation in gastric cancer. Our present systematic review and meta-analysis aims to evaluate the associations between H&E-based TIL density and clinicopathological features and the influence of TILs on prognosis in patients with gastric cancer.

METHODS
This study was performed based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.\textsuperscript{26} Two independent investigators (CT and DS) conducted literature retrieval, selected relevant studies and extracted eligible data. Any disagreement was discussed to reach a consensus.

Search strategy
We conducted a systematic and comprehensive literature retrieval in the following databases: Cochrane Library (1993 to 25 February 2020), PubMed (1966 to 25 February 2020) and Embase (1974 to 25 February 2020). According to MeSH terms, we used the following terms: (neoplasm, stomach or stomach neoplasm or cancer, stomach or cancer of stomach or stomach cancer or cancer of the stomach or gastric neoplasm or neoplasm, gastric or gastric cancer or cancer, gastric or gastric cancer, familial diffuse), in combination with (lymphocytes, tumor infiltrating or tumor-infiltrating lymphocytes or tumor infiltrating lymphocytes or tumor-derived activated cells or tumor derived activated cells or activated cells, tumor-derived). The full electronic search strategy for PubMed can be found in online supplemental tables S1 and S2. We also manually checked reference lists of relevant studies for any other articles.

Inclusion and exclusion criteria
The criteria for inclusion were as follows: (1) studies evaluating the correlation of TILs and either prognosis or clinicopathological characteristics in gastric cancer; (2) TILs were assessed with H&E-stained sections; (3) the number of patients included in each study was more than 50 cases; and (4) studies with sufficient information to determine an estimate of relative risk (RR) or HR and their 95% CIs. The following studies were excluded: (1) duplicates, reviews, letters, comments, case reports, animal trials, cell line trials, conference abstracts and meta-analyses; (2) non-English languages; and (3) lacking information to calculate RR or HR.

Data extraction
The following information was obtained from each study: last name of the first author, publication year, country, sample size, duration of follow-up, cut-off values of TILs, TIL location, number of patients with high or low TILs, gender, age, Lauren classification, histological grade, depth of invasion (T stage), lymph node involvement (N stage), TNM stage, HR and 95% CIs.

Quality assessment
The criteria of Newcastle-Ottawa Scale (NOS) assessment scale\textsuperscript{27} were used to evaluate the quality of eligible studies. The quality score of NOS ranged from 0 to 9. Studies with a score of 6 or more were regarded as high quality.

Statistical analysis
The present meta-analysis was performed through STATA V.12.0. The associations of TILs with clinicopathological characteristics were assessed using pooled RRs with 95% CIs. The relevant HRs and 95% CIs extracted from the original studies were combined into a pooled HR. If HRs and their 95% CIs were not reported directly, they would be calculated from Kaplan-Meier survival curves by the method previously reported by Tierney et al.\textsuperscript{28} The relationship between TILs with prognosis was evaluated using pooled HRs and 95% CIs.

We calculated all effects on the basis of random-effect models. The \( \chi^2 \) test and I\(^2 \) test were used to assess heterogeneity between the enrolled publications. Heterogeneity was considered high for \( p<0.15 \) and I\(^2 \) >50\%.\textsuperscript{29} We performed subgroup analyses according to TIL locations because the function of TILs may vary across different locations. Egger’s test\textsuperscript{30} and Begg’s test\textsuperscript{31} were conducted to detect publication bias. The trim and fill method was used to evaluate the effects of publication bias on the pooled results.\textsuperscript{32} The source of heterogeneity and the stability of the results were assessed by sensitivity analyses.
enrolled in this study were patients with gastric adenocarcinoma. The total number of cases was 2835. All cases from China, two from Japan and one from the USA. All studies were performed in a single-centre setting. The total number of cases was 2835. All cases enrolled in this study were patients with gastric adenocarcinoma. Seven studies (n=2613) reported data on OS.34 36–41 Two studies about animals and cell lines (n=164); Non-gastric cancer (n=112). Did not evaluate on H&E stained sections (n=529).

Six studies36–41 provided HRs, while one study from Kim et al34 only provided Kaplan-Meier survival curves. The HRs and their 95% CIs obtained from Kaplan-Meier curves were calculated according to the methods previously reported.28 The relationships between TIL density and clinicopathological characteristics were reported in seven studies.33–39 TILs are divided into three categories according to their locations: intratumoural TILs (iTu-TILs), stromal TILs (str-TILs) and total TILs (TILs at all locations). iTu-TILs and str-TILs were separately evaluated in one study.35 iTu-TILs or str-TILs33 38 were evaluated in four studies. Total TILs were reported in four other studies.34 36 37 39 TIL density was evaluated manually in eight33–39 41 of the nine included studies. The remaining study was evaluated digitally.40 Four studies33 35 36 38 evaluated TILs based on the IIOBWG recommendations and five other studies34 37 39 41 evaluated TILs based on different approaches. Cut-off values of TILs varied across different studies. All cases were subdivided into a high TIL group and a low TIL group. General information on the enrolled publications is summarised in table 1.

Quality assessment
According to the criteria of NOS, all nine studies were considered to be of adequate quality for the meta-analysis (score ≥6 points) (table 1).

TIL density and survival
Significant heterogeneity existed among the seven studies, including OS data (I2=81.4%, p<0.001). Accordingly, a random-effect model was used. As illustrated in figure 2, compared with low TIL density, high TIL density was significantly associated with better OS (HR=0.68, 95% CI 0.52 to 0.87, p=0.003).

TIL density and clinicopathological characteristics
The relationships between TIL density and clinicopathological characteristics are shown in table 2. High TIL density was significantly associated with lower depth of invasion (T3–T4 vs T1–T2: RR=0.58, 95% CI 0.50 to 0.66, p<0.001), less lymph node involvement (presence vs absence: RR=0.68, 95% CI 0.56 to 0.81, p<0.001) and earlier TNM stage (III–IV vs I–II: RR=0.68, 95% CI 0.55 to 0.83, p<0.001). TIL density was not correlated with age, gender, Lauren classification or histological grade.

Furthermore, we carried out subgroup analyses to explore the associations between TILs in different locations and OS, depth of invasion, lymph node involvement and TNM stage.

Subgroup analysis based on TIL locations
The results of subgroup analysis based on TIL locations are shown in figure 3. High TIL density was significantly related to superior OS independent of their locations (iTu-TILs, p=0.005; str-TILs, p=0.031; total TILs, p=0.039) (figure 3A). iTu-TIL density was significantly correlated with lymph node involvement (p=0.022), but not with TNM stage (p=0.097). A borderline significant trend was shown between iTu-TILs and invasion depth (p=0.052). str-TIL density was significantly related with depth of invasion (p<0.001) and showed borderline association with lymph node involvement (p=0.053) and TNM stage (p=0.054). Total TIL density was significantly associated

with depth of invasion (p<0.001), lymph node involvement (p<0.001) and TNM stage (p=0.028) (figure 3B–D).

Publication bias and sensitivity analysis

Studies included in the OS analysis showed no significant publication bias (pBegg’s=0.764, pEgger’s=0.851). There was no significant publication bias among studies included in the clinicopathological characteristics analysis except for histological grade and Lauren classification (pEgger’s=0.038, pEgger’s=0.020, respectively) (table 2). The trim and fill method indicated that studies were unlikely to be missing in the analysis of the association between TILs and histological grade, while two studies were found missing in the analysis of the association between TILs and Lauren classification (figure 4). The effects of TILs on histological grade and Lauren classification were stable and not changed after having adjusted the publication bias.

Sensitivity analyses were performed by omitting each study in turn. The results showed that the exclusion of any single study has no significant effect on the pooled results (figure 5).

DISCUSSION

To our best knowledge, this is the first meta-analysis of the associations between TIL density evaluated by H&E-stained sections and survival outcomes and clinicopathological features in patients with gastric carcinoma. Our meta-analysis, including 2835 cases from nine qualified observational studies, demonstrated that high TIL density was significantly associated with superior OS, lower depth of invasion, less lymph node involvement and earlier TNM stage. These results indicate that TIL density evaluated by H&E-stained sections can provide important information to predict the survival outcomes of patients with gastric cancer.

Immune checkpoint inhibitors (ICIs), which enhance host antitumour immune response by blocking checkpoint proteins,42 are novel therapeutic options for patients with advanced gastric cancer.43–45 Oncologists and researchers show great interest in TILs because lymphocyte infiltration in tumour is a critical step for immunotherapy.46 Our previous study in colon cancer (n=1532) indicated tumours with high TILs had distinct tumour biology and were more likely to present with lower T and N stage, well or moderate histological grade, right-sided primary location, mutated BRAFV600E and deficient mismatch repair (dMMR).47 Patients with high TILs were more likely to achieve response to ICIs, while...
low TILs represent existing but partially exhausted anti-tumour immunity. High TILs were also associated with the efficacy of chemotherapy in gastric cancer. Li et al. found that high CD3+ or FoxP3+ T cell infiltration was associated with better OS in patients with stage II–III gastric cancer treated with FOLFOX (oxaliplatin+leucovorin+fluorouracil). Our meta-analysis indicates high TILs are related to favourable clinicopathological characteristics, including lower invasion depth, less lymph node involvement and earlier TNM stage, which is interpreted to suggest that high TILs can enhance host antitumour immunity and prevent tumour cells from invading and metastasising. The prognostic value of TILs evaluated by H&E-stained sections in gastric cancer remains controversial. In a study including 148 patients, TIL density showed no association with OS. On the contrary, in another study accruing 1033 patients, high-TIL density was an independent prognostic factor of OS. Our meta-analysis provides strong evidence that H&E-based TIL density is a reliable prognostic biomarker in gastric cancer. This result is consistent with previous reports across multiple cancer types. Moreover, our results are in line with the meta-analyses of prognostic value of TILs evaluated by IHC in gastric cancer. Of note, different TIL subsets detected by IHC had different prognostic values. For example, CD3+ TILs and CD8+ TILs had positive effects on survival, whereas FOXP3+ TILs were negatively correlated with survival.

### Table 2: Association between TIL density and clinicopathological features

<table>
<thead>
<tr>
<th>Clinicopathological features</th>
<th>Studies (n)</th>
<th>RR (95% CI)</th>
<th>P value</th>
<th>Heterogeneity</th>
<th>Publication bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>RR (95% CI)</td>
<td>P value</td>
<td>I² (%)</td>
<td>P value</td>
<td>Effect model</td>
<td>P Egger’s</td>
</tr>
<tr>
<td>Age (elder vs younger)</td>
<td>5</td>
<td>0.97 (0.88 to 1.08)</td>
<td>0.617</td>
<td>0</td>
<td>0.728</td>
</tr>
<tr>
<td>Gender (male vs female)</td>
<td>6</td>
<td>0.99 (0.90 to 1.08)</td>
<td>0.75</td>
<td>0</td>
<td>0.96</td>
</tr>
<tr>
<td>Lauren classification</td>
<td>4</td>
<td>1.13 (0.98 to 1.30)</td>
<td>0.097</td>
<td>0</td>
<td>0.552</td>
</tr>
<tr>
<td>Histological grade (poor vs well, moderate)</td>
<td>5</td>
<td>1.00 (0.73 to 1.37)</td>
<td>0.992</td>
<td>87.7</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Depth of invasion (T3–T4 vs T1–T2)</td>
<td>4</td>
<td>0.58 (0.51 to 0.66)</td>
<td>&lt;0.001</td>
<td>0</td>
<td>0.746</td>
</tr>
<tr>
<td>Lymph node involvement (presence vs absence)</td>
<td>6</td>
<td>0.68 (0.56 to 0.81)</td>
<td>&lt;0.001</td>
<td>63.4</td>
<td>0.012</td>
</tr>
<tr>
<td>TNM stage (III–IV vs I–II)</td>
<td>4</td>
<td>0.68 (0.55 to 0.83)</td>
<td>&lt;0.001</td>
<td>61.8</td>
<td>0.033</td>
</tr>
</tbody>
</table>

RR, risk ratio; TILs, tumour-infiltrating lymphocytes; TNM, tumour, node, metastasis.

### Figure 3: Forest plots of the subgroup analysis of the association between TILs in different locations and prognosis and clinicopathological features.

(A) Overall survival; (B) depth of invasion; (C) lymph node involvement; and (D) TNM stage. RR, relative risk; TILs, tumour-infiltrating lymphocytes; TNM, tumour, node, metastasis.
IHC can define the majority of TIL subsets, provide the detailed information about the composition and functional status of TILs, and improve the evaluating accuracy of TILs. However, IHC-based TIL evaluation is usually conducted on representative tissues, not on whole sections, which cannot provide comprehensive information due to the heterogeneity of TILs in different locations. Moreover, the use of IHC requires additional processing and technical standardisation. The implementation of TILs assessed by IHC sections may be difficult in low-income or middle-income areas and is yet to become routine. Although H&E-based evaluations are unable to capture the complexity of TME, pathologists could be educated to score TILs on H&E-stained sections with high inter-personal concordance compared with untrained pathologists. H&E-based TIL assessment is inexpensive and applicable to routine pathology practice and clinical trials, even in resource-poor settings. To date, TILs assessed by H&E-stained sections are recommended by IIOBWG in breast cancer. Despite no consensus for TIL evaluation and its cut-off in gastric cancer, our meta-analysis suggests TIL density evaluated on H&E-stained sections is a reliable tool to predict survival in this cancer type. Future study is needed to explore the optimal evaluation method and cut-off value for TILs.

TILs are divided into two categories based on their locations, iTu-TILs and str-TILs, which have been evaluated separately or together in various studies. Lymphocytes located in tumour nests are iTu-TILs that directly contact cancer cells, while lymphocytes dispersed in the stroma are defined as str-TILs that do not directly interact with cancer cells. str-TILs have been currently recommended...
as a principal parameter in breast carcinoma pathological reports by IIOBWG. To date, the prognostic value of TIL’s locations varies in different reports and there are no recommendations for evaluation of TILs in gastric cancer. Our meta-analysis showed high TILs on H&E slides were associated with better OS independent of TIL locations, suggesting iTu-TILs, str-TILs and total TILs might be reliable prognostic biomarkers for gastric cancer. A meta-analysis reported that high density of CD3+ TILs evaluated in all areas, including cancer nest and cancer stroma, was associated with longer OS in gastric cancer. One study indicated high iTu-TILs were related to prolonged OS in patients with gastric adenocarcinoma. On the contrary, in another study including 120 patients with Epstein-Barr virus (EBV)-positive gastric adenocarcinoma, the prognostic value was only seen in str-TILs, but not in iTu-TILs. EBV-positive gastric cancers displayed gene mutation (PIK3CA), gene amplification (JAK2, CD274 and PDCD1LG2) and DNA hypermethylation. The association of TIL locations with molecular features awaits further study. Future research that explores the biological difference of TILs in different locations is needed to better understand host antitumour immune responses.

There are several limitations to our study. First, we calculated the HR and 95% CI from the Kaplan-Meier chart in one study, due to a lack of primary data from the original report; this may limit the precision of our data. Second, the methodology for evaluating TILs and its cut-off value varied across different studies, which might affect the results of our meta-analysis. Third, most data were from East Asian populations and the prognostic value of TILs among Western populations should be further studied. Fourth, heterogeneity among different studies is inevitable when meta-analysis is conducted. In addition, publication bias in histological grades and Lauren classification was detected and adjusted by the trim and fill method in our meta-analysis. Finally, retrospective data collection caused selection bias.

Our meta-analysis suggests that TIL density evaluated by H&E-stained sections in observational studies is a reliable biomarker to predict prognosis in patients with gastric cancer. Multicentre, prospective studies are needed to determine the optimal evaluation method (IHC or H&E-stained sections or others) and cut-off values for TILs. Future studies should also focus on the biological difference between TILs in different locations.

Contributors DS and CT designed the study, analysed the data, drafted and edited the manuscript. HJ and YC performed the statistical analysis, interpreted the data, edited and provided critical revision to the article. CW, WW and JC edited and provided critical revision to the article. All authors read and approved the final manuscript.

Funding This study was supported by the Natural Science Foundation of Shandong Province of China (grant no. ZR2013HQ027) and the Science and Technology Innovation Development Program of Jinan, Shandong Province, China (grant no. 202019085).

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data sharing not applicable as no datasets generated and/or analysed for this study. All data relevant to the study are included in the article or uploaded as supplementary information.

Supplemental material This content has been supplied by the author(s), it has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non-Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iDs Halyan Jing http://orcid.org/0000-0003-0913-6898 Dan Sha http://orcid.org/0000-0003-3004-8534

REFERENCES
Differences in immune contextures among the gastric cancer subtypes could be detected by a simple, graphical test. This approach is useful for predicting the neoadjuvant chemotherapy response and overall survival in triple-negative invasive breast cancer. Hum Pathol 2018;80:47–54.


