Prognostic Significance of Tumor-infiltrating T-cells in Ovarian Cancer: a Meta-analysis

Wei-Ting Hwang, Ph.D.[a,b,1], **Sarah F. Adams, M.D.**[a,1], **Emin Tahirovic, M.S.**[b], **Ian S. Hagemann, M.D., Ph.D.**[c,2], and **George Coukos, M.D., Ph.D.**[a,*]

[a]Ovarian Cancer Research Center, University of Pennsylvania, Philadelphia, Pennsylvania 19104, U.S.A.

[b]Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, Pennsylvania 19104, U.S.A.

[c]Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, U.S.A.

Abstract

Objective—The presence of T cells within the epithelial component of tumors, as histologic evidence of anti-tumor immunity, has been associated with a survival advantage in multiple studies across diverse patient cohorts. We performed a meta-analysis of studies evaluating the prognostic value of tumor-infiltrating lymphocytes (TIL) on survival among women with ovarian cancer and to investigate factors associated with variations in this effect, including patient characteristics, surgical outcomes, tumor histology, and study protocols.

Method—Published studies that evaluated the association between TIL and patient survival were identified. Descriptive statistics, outcome data, and study quality were extracted from studies that met inclusion criteria. Hazard ratios and 95% confidence intervals were pooled across studies using the random-effects model. Publication bias was investigated using a funnel plot and heterogeneity was assessed with subgroup analysis and I^2 statistics.

Results—Ten suitable studies comprising 1,815 patients with ovarian cancer were analyzed. Our results demonstrate that a lack of intraepithelial TILs is significantly associated with a worse survival among patients (pooled HR: 2.24, 95% CI; 1.71–2.91). Variations in the prognostic value of TIL status based on debulking status, scoring method, and geographic regions were identified.

Conclusions—Intraepithelial TILs are a robust predictor of outcome in ovarian cancer and define a specific class of patients, whose distinct tumor biology should be taken into account in devising appropriate therapeutic strategies.

[*]Corresponding author, George Coukos, M.D., Ph.D., Ovarian Cancer Research Center, University of Pennsylvania, Room 1315, Biomedical Research Building II/III, 421 Curie Blvd, Philadelphia, PA, 19104-6160 USA, Office: 1-215-573-5016, Fax: 1-215-573-5129, gcks@mail.med.upenn.edu.
[1]Contributed equally to this work
[2]Present Address: Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, Missouri 63110, U.S.A.

Conflicts of Interest Statements
None.

Introduction

Epithelial ovarian cancer is a heterogeneous disease, with significant variation in both the presentation and response to therapy. Prognosis is affected by patient factors, such as age or genetic background, as well as tumor characteristics, including stage, grade, histologic subtype, and chemotherapy sensitivity [1,2]. Recent studies have also identified immunologic biomarkers of prognosis, with longer survival times documented among women with histologic evidence of an anti-tumor immune response. Although T cells are present in the stroma of most tumor specimens, a survival advantage has been associated specifically with the presence of T cells in epithelial tumor islets (intraepithelial tumor-infiltrating lymphocytes, TILs) [3]. In addition to correlations with clinical outcome, evidence favoring an active role for TILs in tumor clearance is provided by data demonstrating that these are oligoclonal T cell populations that recognize tumor antigens *ex vivo* and secrete cytokines characteristic of effector cells [4–7]. With the emergence of immunotherapeutic strategies for the treatment of ovarian cancer, it will be important to validate immunologically relevant tumor biomarkers to optimize patient selection for clinical trials and to prospectively track responses to immunotherapeutics [8,9].

Although all studies of patients with ovarian cancer have described a prognostic advantage associated with intraepithelial TILs, differences in the measurement and characterization of TILs have limited the clinical utility of this biomarker. Questions remain as to whether inconsistencies in results derive from differences in study methodology or whether variable outcomes among diverse patient cohorts illustrate underlying biologic or environmental modifiers of anti-tumor immunity. For example, while some reports have quantified all $CD3^+$ T cells as TILs, others have focused specifically on cytotoxic $CD8^+$ T cells. Additionally, the criteria used to score tumors as TIL-positive or TIL–negative have not been consistent across studies. It is also unclear whether associations between TIL status and survival varied according to the standard prognostic factors, such as age, stage, histology, or surgical outcomes.

The objective of this study is to review the prognostic significance of intraepithelial TILs for overall survival across diverse cohorts of women with ovarian cancer using meta-analytical tools. Our secondary objective is to identify patient, tumor, or methodological characteristics that may explain the variations in the published findings.

Methods

We followed guidelines for the design, analysis, and reporting of meta-analyses of observational studies published by the MOOSE group [10].

Search Strategy

Studies published before December 2010 were identified in PubMed using the following search terms: "ovarian cancer" and "TIL" "lymphocytes", and "T cell". There was no language restriction. The references of all publications were reviewed to identify additional relevant studies.

Study Selection

Studies that met the following criteria were included in the meta-analysis: studies must have (1) been published as original articles; (2) evaluated human subjects; (3) investigated CD3 and/or CD8 lymphocytes in ovarian cancer; (4) reported disease-specific or overall survival; and (5) contained the minimum information necessary to estimate the effects (i.e., hazard ratio) and a corresponding measure of uncertainty (i.e., confidence interval, P-values, standard errors or variance). As an additional criterion, when a single population was

reported in multiple reports, only the report with the most complete data was included to avoid duplication.

Data Extraction

Using a predefined form, data on study cohorts, methodology and results were extracted. The author, year of publication, and region where each study was conducted were noted. The collected patient or tumor characteristics including the number of women in each cohort, the duration of follow-up (mean or median), the ages at the time of surgery (mean, median), and the surgical outcomes (optimal or suboptimal cytoreduction), distributions of stage, grade, and histologic subtype, were recorded. Methodology characteristics analyzed included the markers used (CD3 or CD8), scoring protocols to identify TILs. The number or distribution of TIL-positive or TIL-negative cases, and results of univariate and/or multivariate survival analyses (e.g., log rank test, Cox proportional hazards model) were extracted. We did not contact authors for additional data.

Measures

The endpoint used in this meta-analysis is overall survival. In the absence of overall survival data, disease-specific survival was substituted because these two measures are expected to be similar for ovarian cancer patients. For CD3 and/or CD8 TIL, study-defined binary variables indicating either the presence (*versus* absence), positive (*versus* negative), or high (*versus* low) marker expression were used and described as "TIL-positive" or "TIL-negative" for this meta-analysis.

Assessment of Study Quality

Study quality was independently rated by two coauthors (WH, ET). Because there is no validated instrument to measure study quality for prognostic marker studies in an observational setting, we adapted the Newcastle-Ottawa Scale and the framework suggested by Altman [11,12]. Briefly, this instrument assesses the quality of studies based on study population (three criteria), prognostic variables (four criteria), outcome measures (two criteria), study duration (one criteria), and statistical analysis (two criteria). Each of the criteria was rated on a three-level scale; zero (no report or criterion not met), one (criterion partially met), or two (criterion was met). Scores from individual criteria were summed and divided by the maximum possible score to produce a total score between zero and one, where higher scores denote greater study quality. The final quality ratings were based on the averaged score (95% limits of agreement: −0.22, 0.11)

Statistical Analysis

The hazard ratio (HR) was used as a measure of the prognostic value, and defined as the hazard of death for women with TIL-negative tumors over the hazard of death for women with TIL-positive tumors, so that a hazard ratio >1 indicated an elevated risk of death in cases lacking intraepithelial TIL. Following the method described in Parmar *et al.* [13], the log-hazard ratio and its standard error for each study were derived. All but one study reported results of a Cox regression analysis; for the remaining study, the log-hazard ratio and its standard error were estimated indirectly based on the reported P-value for the log rank test and the number of deaths observed in the study. If results of both univariate and multivariate Cox regression analyses were reported, we used estimates from the multivariate Cox regression model for a more direct estimate of the effect of TIL after controlling for potential confounding variables. In two studies where results for both CD3 and CD8 were reported, the estimates based on CD8 markers were used for the primary analysis. To account for heterogeneity among studies, random effects models were used to estimate pooled HRs [14]. The 95% confidence interval (CI) for the pooled HR was reported.

Homogeneity of effects across studies was assessed using I^2 statistics [15]. This statistic describes the percentage of total variation across studies that are due to heterogeneity rather than chance (25% low heterogeneity, 50% medium, 75% high).

In secondary analyses, pooled HRs were estimated by specific TIL markers (CD3, CD8). One study which did not distinguish between CD3 and CD8 was included in the CD8 analysis. Subgroup analyses were carried out to investigate potential sources of between-study heterogeneity and to assess whether conclusions were sensitive to restricting studies to subgroups that might have different prognostic effects. Subgroups were defined according to TIL scoring algorithm (zero versus >0), specimen processing (paraffin-embedded, tissue micro-arrays (TMA), cryosection), debulking status (optimal only versus mixed), histology (serous only versus mixed), stage (III/IV only versus mixed), grade (>75% grade three, versus less), and by geographic region (North America, Europe, Japan). Tests for effects-subgroup interaction were performed. Publication bias was evaluated by inspecting the symmetry of the funnel plot and formally tested with Begg's adjusted rank correlation test [16,17]. Statistical analysis was conducted with Stata version 11 (College Station, Texas) and Review Manager Version 5·0 (The Cochrane Collaboration).

Results

Study Selection

Of 18 potentially eligible articles, ten met the inclusion criteria and were evaluated further. Figure 1 provides a summary of the selection process [3,18–32]. The research quality among the selected studies was high; with median quality score of 0.86 (range 0.75 to 0.92).

Patient Cohorts

Characteristics of patient cohorts from the analyzed studies are shown in Table 1. The median number of women evaluated per study was 142 (range 70 to 500), with a total of 1,815 subjects across all studies. The mean age in all cohorts was similar, ranging from 55–62 years. The percentage of women with stage III–IV disease varied from 16.8% to 100%, with four studies including only advanced cases [3,19,31–32]. Most patients had serous tumors, with two studies focusing exclusively on women with serous cancers [19,31], and 41% to 100% of patients had grade 3 tumors. Surgical outcomes also varied: although the criteria defining optimal cytoreduction were not specified (<2cm, 1cm, or no macroscopic residual disease), in only half of the reports had >60% of women considered optimally debulked by study authors. Nine studies reported the duration of follow-up, with the median duration ranging from 1.8 to 10 years. Eight studies reported analysis results for overall survival while two used cause-specific survival as the primary endpoint.

Quantification of intraepithelial tumor-infiltrating lymphocytes

All studies scored intraepithelial TIL based on immunohistochemical analysis of tumor specimens: two used cryosections, five used paraffin-embedded tissue, and three used tissue microarrays (TMA). Results are summarized in Table 2. The definition for positive staining varied, ranging from 1 to 10 cells per 200X high-power field (HPF). Four studies used percentiles (e.g. median, tertile) to determine TIL positivity [19,24,26,28]. Among six studies reporting intraepithelial CD3+ TIL, the proportion of patients with positive scores ranged from 39% to 90.2% with a median of 60%. For intraepithelial CD8+ TIL, the proportion of patients with positive scores ranged from 19.4% to 81.4%, with a median of 62%. Studies reporting a higher proportion of intraepithelial TILs also reported higher proportions of subjects with optimal debulking (Spearman correlations: 0.81, 0.31 for CD3+ and CD8+, respectively), and lower proportions of subjects with serious histology (Spearman correlations: −0.80, −0.08), more subjects with grade 3 (Spearman correlations:

−0.41, −0.22), and stage III/VI (Spearman correlations: −0.73, −0.18) tumors. None of the correlations reached statistical significance at 0.05 level because of small number of studies analyzed.

Prognostic value of intraepithelial CD3[+] and/or CD8[+] lymphocytes

All studies reported that women with intraepithelial TIL positive tumors (CD3[+] or CD8[+]) demonstrated significantly longer overall survival (HRs ranging from 1.33 to 7.62; all P-values <0.05), with half greater than or equal to 2.60. Results of multivariate analyses were used for eight studies. In the overall meta-analysis, the HRs for CD8[+] TILs were used for seven studies, CD3[+] TILs for two, and one study did not specify which marker was used [27]. A meta-analysis of all studies confirms a significant association between intraepithelial TILs (CD3[+] or CD8[+]) and survival, with a pooled HR of 2.24 for TIL-negative cases (95% CI; 1.71–2.91). A moderate to high heterogeneity between studies was noted (I^2=68%). A forest plot (Figure 2) demonstrated that these studies are in agreement for the reported hazard ratios. For CD3[+] TIL, we found a pooled HR of 1.74 (95% CI; 1.14–2.66) using 5 studies, indicating a significantly increased risk for death for those lacking intraepithelial CD3[+] T cells. For CD8[+] TIL, a pooled HR of 2.19 (95% CI; 1.60–2.98) for TIL-negative cases from 8 studies was obtained. Heterogeneity between studies remained substantial with I^2 statistics of 79% and 70% for CD3 and CD8 markers respectively. In the three studies where results for both CD3 and CD8 were reported [24,30,32], the reported HRs showed the same direction of association (all >1) but were less significant for the CD3 markers.

Subgroup analysis on the effects of intraepithelial TIL on survival by study, patient or tumor characteristics

In a stratified analysis, intraepithelial TIL-positive status was continued to be associated with a reduced hazard of death (Table 3), but the prognostic significance varied based on surgical outcomes, scoring method, and geographic region. One study exclusively evaluating optimally debulked cases [30] reported a significantly lower HR than studies including both optimally and suboptimally cytoreduced patients (pooled HRs: 1.33 *versus* 2.38). Three studies using a complete absence (0) of T cells to define TIL-negative cases [3,27,30–31] reported a pooled HR of 1.53 (95% CI; 1.22–1.93), which is significantly lower than the pooled HR of 2.67 (95% CI; 2.02–3.53) for studies with higher cut-off values. In addition, results varied regionally, with smaller effects among cohorts from North America (pooled HR: 1.68), and a large HR of 7.62 (95% CI; 2.76–21.08) in a single Japanese study [25]. No difference was seen with regards to tumor characteristics, with similar HRs between groups defined by distributions of stage, grade or histologic subtype. No significant difference in the pooled HRs was noted based on the use of CD3 or CD8 expression, or differences in specimen processing.

Begg's rank-correlation test indicated no publication bias at the 0.05 level (P=0.09). Examining the funnel plot (Figure 3) however demonstrated evidence of asymmetry around the value of the pooled HR, but mostly due to one study with both large estimated hazard ratio and standard error [25]. The statistical evidence for publication bias was reduced if this study was removed (Begg's test, P=0.21)

Discussion

This meta-analysis confirms that the intraepithelial TILs in ovarian cancer specimens are a robust biomarker for overall survival of women with this disease. This is true in spite of the significant heterogeneity among studies regarding both patient characteristics and scoring methodology. The majority of studies in this meta-analysis used the CD8 marker to specifically evaluate cytotoxic T cells. Although no significant difference was seen in the

pooled HRs based on analysis of CD3 or CD8 expression, intraepithelial CD8 TIL showed a more consistent and stronger association with survival than CD3 TIL. Thus, CD8 staining should be used as the standard for evaluation of intraepithelial TIL in ovarian cancer specimens. Further, a significant difference was seen in the HRs based on scoring method used to evaluate TIL. While TIL scores represent an underlying continuous variable, a standardized measure of TIL positivity would facilitate future studies. Because significantly larger HRs were noted in studies that used greater than zero cut-offs (e.g., 3–10 cells/HPF) for a positive score, and 5 cells/HPF approximately represents the midpoint of those cut-off values, we propose that >5 CD8$^+$ cells/200X HPF should be defined "TIL-positive" in ovarian tumors.

The strong association of intraepithelial TILs with better outcomes strongly suggests that antitumor effector mechanisms are spontaneously activated in a proportion of patients with advanced ovarian cancer and have a significant impact on tumor growth and metastasis. Indeed, tumor-reactive antibodies and T cells have been isolated from peripheral blood of patients with advanced stage ovarian cancer at diagnosis [12, 13], while oligoclonal tumor-reactive T cells have been isolated from ascites or tumors [18–22]. Definitive studies are needed to dissect the potential mechanisms through which effective TILs can control tumor growth, but direct cytotoxic effect on tumor cells as well as a potent antiangiogenic effect of angiostatic cytokines such as interferon-gamma, and chemokines such as interferon-inducible CXCL9 and CXCL10, all of which are elevated in tumors with intraepithelial TILs [3], could contribute to reduce tumor growth and prolong survival.

In spite of numerous publications demonstrating the impact of TILs in ovarian cancer, this biomarker is still not used clinically in treatment planning or in evaluating the prognosis of individual patients. It is possible that the differences in the definition of TILs as well as variations in study methodology have impeded the adoption of TIL status by pathologists or physicians responsible for the treatment of women with ovarian cancer. One goal of this report was to consolidate the data from these disparate studies to facilitate the use of this prognostic marker in clinical practice. Furthermore, as our understanding of the possible interactions between immune mechanisms and chemotherapy increases, TILs could be used to guide selection of drugs in the future [9]. For example, pegylated liposomal doxorubicin as well as phase-specific drugs like taxanes or topotecan, but not time-specific drugs like gemcitabine, were found to interact positively with immune activation in a mouse model of ovarian cancer [36–37], suggesting that patients with intraepithelial TILs might benefit more from specific drug combinations. In addition, TILs could be developed as a potential predictive biomarker for immunotherapy, given the increasing opportunities for immunotherapeutic drugs available for testing. These clinical opportunities are discussed in detail elsewhere [8–9].

Interestingly, this analysis also demonstrates that the prognostic value of TILs may vary depending on surgical outcomes. Specifically, protective effects associated with TILs were significantly reduced in the single study that included only optimally debulked cases compared to studies that also include suboptimally debulked cases. Although surgical debulking has been shown to be an independent prognostic factor in studies of TIL, this finding is consistent with the interaction between TIL status and surgical outcomes described previously. One study reported that the benefit of surgical debulking was most pronounced among women who lacked TILs, suggesting that women with spontaneous anti-tumor immunity may be able to control small volume disease following suboptimal debulking and cytotoxic therapy, while women lacking this response require maximal surgical effort to improve outcomes [32]. Interestingly, Zhang *et al.* [3] reported that TIL-positive women were more likely to have been optimally cytoreduced, suggesting that anti-tumor immune responses may have restricted disease spread or altered tumor biology in a

manner that facilitated surgical resection. Whether TILs can be used to direct surgical decision-making requires further study.

An interesting finding of this study relates to the regional differences in results, with greater prognostic effects seen among studies done in Japan and Europe than in North America. This may result from differences in surgical technique, with geographic region serving as a surrogate marker for the proportion of optimally cytoreduced patients where more aggressive debulking practices reported in the United States [33–35]. However, large proportions of women were considered optimally cytoreduced in the European studies included in this analysis. Whether this represents differences in the criteria used to determine optimal cytoreduction cannot be determined with the available data. Other possibilities that might explain regional variation include immune-modifying factors, such as dietary practices, rates of obesity, other environmental variables, or genetic differences (e.g., polymorphisms in genes affecting immune mechanisms), and possible differences in access to healthcare resources. While further study is needed to identify genetic, behavioral or environmental modifiers for the effects of TILs, it is notable that across diverse patient cohorts, a survival advantage is consistently demonstrated among women with intraepithelial TIL.

Certain limitations of our analysis must be acknowledged. Only ten studies were suitable for this meta-analysis, and all of them were retrospective in nature. While the inclusion of studies evaluating different lymphocyte populations (CD3 or CD8) and different scoring criteria limits the direct comparison of patients across cohorts, it also supports the conclusion that TIL score is a robust prognostic marker. This study did not evaluate the effects of other lymphocyte populations such as FoxP3[+] regulatory T cells, which may modulate the CD8+ effector response and thus impact the prognostic significance of TIL. Finally, most studies lacked information regarding progression-free survival or response to adjuvant therapy, and prospective studies will be required to determine the interactions of TIL status with treatment drugs.

In conclusion, we have demonstrated that intraepithelial CD3 and CD8 TIL are useful immunologic biomarkers in predicting survival and that CD8 is more robust for stratifying women with ovarian cancer based on prognosis. In our study, the prognostic value of TILs persists among populations with diverse histologic tumor subtypes, regardless of the stage or grade of disease. Regional variations had suggested that other modifiable factors to anti-tumor immunity may exist. Finally, the HR associated with TIL status may be highest among suboptimally debulked patients but whether this marker can be used to direct surgical decision-making requires further study. With the recent development of immunotherapeutic strategies for the treatment of advanced or recurrent ovarian cancer, and evidence of the immunomodulatory effects of standard chemotherapeutic agents, the use of immune-based biomarkers takes on added importance. If validated, TIL scores could be used not only to select women for clinical trials of cancer vaccines or adoptive T cell therapies, but could also be used to select combinations of chemotherapeutic drugs that interact positively with antitumor immune mechanisms.

Highlights

- Meta-analysis of ten studies with 1,815 ovarian cancer patients confirms a significant survival advantage associated with tumor-infiltrating lymphocytes (TIL).

- This effect was demonstrated across diverse patient cohorts regardless of the tumor grade, stage, or histologic subtype studied.

> • A standardized measure of CD8$^+$ TILs in ovarian cancer will facilitate the clinical use of this robust biomarker.

Acknowledgments

References

1. Winter WE, Maxwell GL, Tian C, et al. Prognostic factors for stage III epithelial ovarian cancer: a Gynecologic Oncology Group Study. J Clin Oncol. 2007; 25:3621–3627. [PubMed: 17704411]

2. Nguyen TT, Wright JD, Powell MA, et al. Prognostic factors associated with response in platinum retreatment of platinum-resistant ovarian cancer. Int J Gynecol Cancer. 2008; 18:1194–1199. [PubMed: 18217964]

3. Zhang L, Conejo-Garcia JR, Katsaros D, et al. Intratumoral T cells, recurrence, and survival in epithelial ovarian cancer. N Engl J Med. 2003; 348:203–213. [PubMed: 12529460]

4. Hayashi K, Yonamine K, Masuko-Hongo K, et al. Clonal expansion of T cells that are specific for autologous ovarian tumor among tumor-infiltrating T cells in humans. Gynecol Oncol. 1999; 74:86–92. [PubMed: 10385556]

5. Halapi E, Yamamoto Y, Juhlin C, et al. Restricted T cell receptor V-beta and J-beta usage in T cells from interleukin-2-cultured lymphocytes of ovarian and renal carcinomas. Cancer Immunol Immunother. 1993; 36:191–197. [PubMed: 8439980]

6. Kooi S, Freedman RS, Rodriguez-Villanueva J, et al. Cytokine production by T-cell lines derived from tumor-infiltrating lymphocytes from patients with ovarian carcinoma: Tumor-specific immune responses and inhibition of antigen-independent cytokine production by ovarian tumor cells. Lymphokine Cytokine Res. 1993; 12:429–437. [PubMed: 8123759]

7. Dadmarz RD, Ordoubadi A, Mixon A, et al. Tumor-infiltrating lymphocytes from human ovarian cancer patients recognize autologous tumor in an MHC class II-restricted fashion. Cancer J Sci Am. 1996; 2:263. [PubMed: 9166543]

8. Kandalaft LE, Powell DJ Jr, Singh N, Coukos G. Immunotherapy for Ovarian Cancer: What's Next? J Clin Oncol. 2011; 29:925–933. [PubMed: 21079136]

9. Kandalaft LE, Singh N, Liao JB, et al. The emergence of immunomodulation: Combinatorial psychotherapy opportunities for the next decade. Gynecol Oncol. 2010; 116:222–233. [PubMed: 19959212]

10. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting Meta-analysis of observational studies in Epidemiology (MOOSE) group. JAMA. 2000; 283:2008–2012. [PubMed: 10789670]

11. Wells, GA.; Shea, B.; O'Connell, D.; Peterson, J.; Welch, V.; Losos, M.; Tugwell, P. [accessed July 24, 2011] The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

12. Altman DG. Systematic reviews of evaluations of prognostic variables. Br Med J. 2001; 323:224–228. [PubMed: 11473921]

13. Parmar MKB, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. Stat Med. 1998; 17:2815–2834. [PubMed: 9921604]

14. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986; 7:177–188. [PubMed: 3802833]

15. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002; 21:1539–1558. [PubMed: 12111919]

16. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. Br Med J. 1997; 315:629–634. [PubMed: 9310563]

17. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics. 1994; 50:1088–1101. [PubMed: 7786990]

18. Nelson BH. The impact of T-cell immunity on ovarian cancer outcomes. Immunol Rev. 2008; 222:101–116. [PubMed: 18363996]

19. Callahan MJ, Nagymanyoki Z, Bonome T, et al. Increased HLA-DMB Expression in the Tumor Epithelium Is Associated with Increased CTL Infiltration and Improved Prognosis in Advanced-Stage Serous Ovarian Cancer. Clin Cancer Res. 2008; 14:7667–7673. [PubMed: 19047092]

20. Tsiatas ML, Gyftaki R, Liacos C, et al. Study of T lymphocytes infiltrating peritoneal metastases in advanced ovarian cancer: associations with vascular endothelial growth factor levels and prognosis in patients receiving platinum-based chemotherapy. Int J Gynecol Cancer. 2009; 19:1329–1334. [PubMed: 20009885]

21. Dogan Y, Erdogan G, Pestereli HE, Tirak B, Karaveli S. Effect of intratumoral lymphocyte density on the disease-free survival and overall survival in malignant epithelial tumors of the ovary. Eur J Gynaecol Oncol. 2009; 30:523–526. [PubMed: 19899407]

22. Shah CA, Allison KH, Garcia RL, Gray HJ, Goff BA, Swisher EM. Intratumoral T cells, tumor-associated macrophages, and regulatory T cells: association with p53 mutations, circulating tumor DNA and survival in women with ovarian cancer. Gynecol Oncol. 2008; 109:215–219. [PubMed: 18314181]

23. Milne K, Kobel M, Kalloger SE, et al. Systematic analysis of immune infiltrates in high-grade serous ovarian cancer reveals CD20, FoxP3 and TIA-1 as positive prognostic factors. PLoS One. 2009; 4:e6412. [PubMed: 19641607]

24. Sato E, Olson SH, Ahn J, et al. Intraepithelial CD8+ tumor-infiltrating lymphocytes and a high CD8+/regulatory T cell ratio are associated with favorable prognosis in ovarian cancer. Proc Natl Acad Sci U S A. 2005; 102:18538–18543. [PubMed: 16344461]

25. Hamanishi J, Mandai M, Iwasaki M, et al. Programmed cell death 1 ligand 1 and tumor-infiltrating CD8+ T lymphocytes are prognostic factors of human ovarian cancer. Proc Natl Acad Sci U S A. 2007; 104:3360–3365. [PubMed: 17360651]

26. Tomsova M, Melichar B, Sedlakova I, Steiner I. Prognostic significance of CD3+ tumor-infiltrating lymphocytes in ovarian carcinoma. Gynecol Oncol. 2008; 108:415–420. [PubMed: 18037158]

27. Han LY, Fletcher MD, Urbauer DL, et al. HLA class I antigen processing machinery component expression and intratumoral T-Cell infiltrate as independent prognostic markers in ovarian carcinoma. Clin Cancer Res. 2008; 14:3372–3379. [PubMed: 18519766]

28. Leffers N, Gooden MJ, de Jong RA, et al. Prognostic significance of tumor-infiltrating T-lymphocytes in primary and metastatic lesions of advanced stage ovarian cancer. Cancer Immunol Immunother. 2009; 58:449–459. [PubMed: 18791714]

29. Milne K, Barnes RO, Girardin A, et al. Tumor-infiltrating T cells correlate with NY-ESO-1-specific autoantibodies in ovarian cancer. PLoS One. 2008; 3:e3409. [PubMed: 18923710]

30. Clarke B, Tinker AV, Lee CH, et al. Intraepithelial T cells and prognosis in ovarian carcinoma: novel associations with stage, tumor type, and BRCA1 loss. Mod Pathol. 2009; 22:393–402. [PubMed: 19060844]

31. Stumpf M, Hasenburg A, Riener MO, et al. Intraepithelial CD8-positive T lymphocytes predict survival for patients with serous stage III ovarian carcinomas: relevance of clonal selection of T lymphocytes. Br J Cancer. 2009; 101:1513–1521. [PubMed: 19861998]

32. Adams SF, Levine DA, Cadungog MG, et al. Intraepithelial T cells and tumor proliferation: impact on the benefit from surgical cytoreduction in advanced serous ovarian cancer. Cancer. 2009; 115:2891–2902. [PubMed: 19472394]

33. Chi DC, Eisenhauer EL, Lang J, et al. What is the optimal goal of primary cytoreductive surgery for bulky stage IIIc epithelial ovarian carcinoma (EOC)? Gynecol Oncol. 2006; 103:559–564. [PubMed: 16714056]

34. Eisenhauer EL, Abu-Rustum NR, Sonoda Y, et al. The effect of maximal surgical cytoreduction on sensitivity to platinum-taxane chemotherapy and subsequent survival in patients with advanced ovarian cancer. Gynecol Oncol. 2008; 108:276–281. [PubMed: 18063020]

35. Hoskins WJ, McGuire WP, Brady MF, et al. The effect of diameter of largest residual disease on survival after primary cytoreductive surgery in patients with suboptimal residual epithelial ovarian carcinoma. Am J Obstet Gynecol. 1994; 170:974–980. [PubMed: 8166218]

36. Alagkiozidis I, Facciabene A, Carpenito C, et al. Increased immunogenicity of surviving tumor cells enables cooperation between liposomal doxorubicin and IL-18. J Transl Med. 2009; 7:104. [PubMed: 20003308]

37. Alagkiozidis I, Facciabene A, Tsiatas M, et al. Time-dependent cytotoxic drugs selectively cooperate with IL-18 for cancer chemo-immunotherapy. J Transl Med. 2011; 9:77. [PubMed: 21609494]

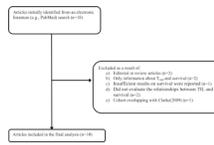**Figure 1.**
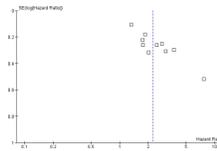Flowchart of study identification and inclusion

Figure 2.
Forest plot of associations of TILs with overall survival in select studies

Figure 3.
Funnel plot for assessing publication bias in select studies.

Table 1

Characteristics of included studies

Author (year) [ref]	No.	Country	Mean age (SD or range)	Serous Histology (%)	Optimal debulking (%)	Stage III–IV (%)	Grade 3 (%)	Median follow-up (yr)	Quality score [c]
Zhang (2003) [3][a]	174	Italy	58.3 (26–80)	77	51	100	67.5	2.00	0.90
Sato (2005) [24]	117	USA	62* (33–89)	78	47	86	90	2.55	0.85
Hamanishi (2007) [25]	70	Japan	55 (11.18)	40	70	55.7	NR	5.19**	0.83
Han (2008) [27]	150	USA	61* (28–89)	78.7	60	91.3	92[b]	1.85	0.79
Callahan (2008) [19]	184	USA	59.9* (20–95)	100	80	100	100	NR	0.88
Tomsova (2008) [26]	116	Czech R.	55 (27–82)	47	52	74	41	3.25	0.75
Adams (2009) [32]	134	USA	62 (26–84)	94	45	100	79.9	3.66**	0.90
Clarke (2009) [30]	500	Canada	58.1 (25–89)	42.3	100	16.8	57.2	5.30	0.89
Leffers (2009) [28]	270	Netherlands	56.8 (13·8)	54.4	58.1	65.6	41.9	3.12**	0.92
Stumpf (2009) [31]	100	Germany	64 (32–85)	100	74	100	68.0	1.87	0.75

*
median age.

**
mean follow-up. NR: not reported.

[a]The statistics was based on a cohort of 74 patients with complete response. The median follow-up was inferred from the no. at-risk table for overall survival.

[b]Grade 2 and 3 combined.

[c]Ranges from 0 to 1 with higher scores indicate better study quality.

Table 2

Distributions of TIL status in studies

Author (year)	CD3		CD8		Definition of positive TIL	Specimen processing
	TIL + N(%)	TIL – N(%)	TIL + N(%)	TIL – N(%)		
Zhang (2003)	102(58.6)	72(41.4)			Any cells/HPF	Cryosections
Sato (2005)	78 (66.7)	39 (33.3)	78(66.7)	39(33.3)	Upper two tertiles (>3.3 cells/20xHPF)	Paraffin-embedded
Hamanishi (2007)			32(45.7)	38(54.3)	$\geq$5 cells/0.0625mm^2	Paraffin-embedded
Han (2008) [a]			106 (70.6)	44(29.4)	Any cells/HPF	Paraffin-embedded
Callahan (2008)			46 (25.0)	138 (75.0)	Top quartile	Paraffin-embedded
Tomsova (2008)	58 (50.0)	58(50.0)			>Median (125 cells/mm^2)	Paraffin-embedded
Adams (2009)	52 (39.0)	82 (61.0)	28 (19.4)	106(80.6)	>10 cells/HPF	Cryosections
Clarke (2009) [b]	296(60.4)	194(29.6)	278(57.2)	208(42.8)	Any cells/2 0.6mm cores	TMA
Leffers (2009)			151 (66.5)	76(33.5)	Upper two tertiles as number of cells/mm^2	TMA of primary tumors
Stumpf (2009)	83 (90.2)	9 (9.8)	71 (81.4)	26 (18.6)	Any cells/HPF	TMA of primary tumors

[a] Did not distinguish for which cell type and listed as CD3/CD8.

[b] Numbers of subjects were inferred from reported percentages and total cohort size.

HPF: 200X high power field.

Table 3

Pooled hazard ratio for overall survival and heterogeneity analysis

	No. of study	HR (95% CI)	Overall effects P-value	Subgroup difference P-value [a]	I² statistics
All studies	**10**	**2.24 [1.71, 2.91]**	**< 0.00001**		**69%**
TILs [b]					
CD3	5	1.74 [1.14, 2.66]	0.01	0.39	79%
CD8	8	2.19 [1.60, 2.98]	< 0.00001		70%
TIL scoring					
0 cut - off	3	1.53 [1.22, 1.93]	0.0003	0.003	33%
>0 cut - off	7	2.67 [2.02, 3.53]	< 0.00001		41%
Specimen					
TMA	3	2.00 [1.18, 3.39]	0.01		81%
Paraffin	5	2.72 [1.75, 4.23]	< 0.00001	0.40	65%
Cryosection	2	1.88 [1.39, 2.55]	0.0001		0%
Debulking status					
Optimal only	1	1.33 [1.08, 1.63]	0.008	0.0002	NA
Mixed	9	2.38 [1.89, 3.00]	< 0.00001		42%
Histology					
Serous only	2	2.01 [1.43, 2.81]	< 0.0001	0.53	4%
Mixed	8	2.33 [1.67, 3.25]	< 0.00001		75%
Stage					
III and IV only	4	1.94 [1.55, 2.42]	0.00001	0.27	0%
Mixed	6	2.60 [1.62, 4.18]	< 0.0001		82%
Grade					
> 75% grade 3	3	2.09 [1.51, 2.88]	< 0.00001	0.98	10%
≤ 75% grade 3	6	2.08 [1.51, 2.86]	< 0.00001		73%
Region					
North America	5	1.68 [1.30, 2.16]	< 0.0001	0.006	38%
Europe	4	2.47 [1.84, 3.31]	< 0.00001		35%
Japan	1	7.62 [2.76, 21.04]	< 0.0001		NA

NA: not applicable.

[a] Test homogeneity between subgroup.

[b] Three studies reported survival analysis for both CD3 and CD8, Han (2008) did not distinguish CD3/CD8 and was listed under CD8 in the current analysis.