High CD3D/CD4 ratio predicts better survival in muscle-invasive bladder cancer

Ming-Jun Shi1,2
Xiang-Yu Meng3,4
Qiu-Ji Wu5
Xiong-Hui Zhou3

1Institut Curie, PSL Research University, CNRS, UMR 144, Paris F-75005, France;
2Department of Urology, Beijing Friendship Hospital, Capital Medical University, Beijing, 430071, People’s Republic of China;
3Hubei Key Laboratory of Agricultural Bioinformatics, College of Informatics, Huazhong Agricultural University, Wuhan 430070, People’s Republic of China;
4Department of Urology, Zhongnan Hospital of Wuhan University, Wuhan, 430071, People’s Republic of China;
5Department of Oncology, Zhongnan Hospital of Wuhan University, Wuhan University, 430071, People’s Republic of China

Correspondence: Qiu-Ji Wu
Department of Oncology, Zhongnan Hospital of Wuhan University, 169 Donghu Road, Wuhan 430071, People’s Republic of China
Tel +861 887 228 4543
Email wuqiuji@126.com

Xiang-Yu Meng
Department of Urology, Zhongnan Hospital of Wuhan University, 169 Donghu Road, Wuhan 430071, People’s Republic of China
Tel +861 582 726 8359
Email mengxy_whu@163.com

Background: Bladder cancer is a common malignancy that affects the human urinary tract. Muscle-invasive bladder cancer (MIBC) is aggressive and has poor prognosis. Previous studies have reported that the tumor-infiltrating lymphocytes (TILs) were associated with MIBC outcome; however, inconsistency remains and mRNA level TIL markers’ prognostic significance in MIBC is unclear.

Materials and methods: In the present study, we reanalyzed data from four public datasets (the Cancer Genome Atlas for investigation; and CIT, GSE5287, and GSE31684 for validation) to examine the prognostic significance of CD3D, CD4, CD8A, CD3D/CD4 and CD3D/CD8A in MIBC.

Results: We found that the CD3D/CD4 ratio was a stable independent prognostic factor in MIBC (beta = −0.87, P = 0.025); high CD3D/CD4 ratio predicted better survival in MIBC, and the power of this association was much stronger in basal-squamous tumors (beta = −4.73, P = 2.6E-06). We also noted that the CD4 expression was significantly higher than CD3D (P < 0.05), indicating the presence of CD3+CD4+ cells which could be immune-suppressing.

Conclusion: The CD3D/CD4 ratio can be viewed as a prognostic marker and a rough measurement for the interaction between immune-effecting CD3+ TILs and immune-suppressing CD3+CD4+ cells in MIBC, and this interaction may play a particularly important role in anti-cancer immunity in basal-squamous tumors as it has a very strong association with survival in this subtype, and may be used to select potential responders to immunotherapy.

Keywords: bladder cancer, muscle-invasive, tumor-infiltrating lymphocytes, CD3, CD4, CD8, prognosis, immunotherapy, basal-squamous subtype

Introduction
Bladder cancer (BC) is the most common malignancy affecting the human urinary tract, with approximately 430,000 new cases diagnosed in 2012 and 165,000 annual deaths.1 Muscle-invasive bladder carcinoma (MIBC, defined as BC with pathologic T≥2) is a major clinical issue and displays unfavorable prognosis, with 5-year survival rate being 50%. Radical cystectomy in combination with neo-adjuvant and/or adjuvant chemoradiotherapy is currently recommended treatment option,2,3 and synchronous radiotherapy with chemotherapy4 or bladder-sparing trimodal strategies5 are also alternatives to cystectomy in properly selected patients for MIBC. However, no major improvement in MIBC patients’ survival has been achieved over the last 20 years.6 Quite recently, the emerging immune checkpoint inhibitors have yielded promising efficacy and improvement on overall survival (OS) in subjects with
advanced and metastatic BC,⁷,⁸ but there is yet no established protocol to select these responders.

Theoretically, durable response induced by anti-cancer immunotherapy depends on the following elements: 1) tumor immunogenicity, mainly associated with tumor mutation burden (TMB) and neo-antigen load; 2) tumor microenvironment, in particular the abundance and composition of tumor-infiltrating lymphocytes (TILs); and 3) the effects of tumor-associated immune suppression.⁹,¹⁰ BC is known as the third top cancer in terms of TMB, just following melanoma and non-small-cell lung cancer (NSCLC),¹¹,¹² therefore, it is considered a possibly ideal target disease of immunotherapy.⁸,¹³,¹⁴

Meanwhile, identification of TIL markers able to predict prognosis will largely improve risk-evaluation and therapeutic decision-making for BC, especially MIBC patients. Previously, several articles reported that TILs and their population structure were associated with prognosis in BC patients,¹⁵⁻¹⁷ whereas others found negative prognostic value of TILs in a global pan-cancer analysis.¹⁸ Debates and uncertainty remain for this issue. Moreover, traditional investigations regarding the prognostic value of TILs were based on either immunohistochemistry or flow cytometry counting of infiltrated immune cells, but both are time-consuming and require considerable resources. Gene expression biomarkers with adequate predictability can be an alternative.

Recently, the increasing accessibility of cancer-omics datasets have largely boosted the discovery of cancer biomarkers. Refined data mining of bioinformatics datasets, especially those stored in open data repositories like the Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO), are becoming the very first step of the translational pipeline of identifying clinically useful biomarkers.

We noted that the association between mRNA expression of key TIL markers (CD3, CD4, and CD8A) and clinical outcome is not clear in BC. In the present study, we aimed to investigate the prognostic significance of CD3D (total T cells), CD4 (helper T cells) and CD8A (cytotoxic T cells) gene expression level and the ratios of CD3D/CD4 and CD3D/CD8A expression in MIBC patients using public datasets (one investigation cohort from TCGA,¹⁹ and three validation cohorts from GEO).

**Materials and methods**

**Datasets**

The TCGA dataset (investigation dataset), including the transcriptome data by RNA-sequencing and clinical data of a cohort of 407 MIBC patients, was downloaded from the Genomic Data Commons data portal (https://gdc.cancer.gov). The information of these patients’ mRNA clustering subtype was obtained from the supplementary materials of Robertson et al.¹⁹ The validation datasets (CIT, GSE5287, and GSE31684) were downloaded from the GEO data portal (https://www.ncbi.nlm.nih.gov/geo) and supplementary materials of corresponding publications. The characteristics of the included datasets were summarized in Table S1.

**Statistical analysis**

For the TCGA dataset, the RNA-Seq by Expectation-Maximization expression values of CD3D, CD4, and CD8A were subjected to log2 transformation. The association between CD3D, CD4, and CD8A expression, ratios of CD3D/CD4 and CD3D/CD8A, and clinical characteristics including age, gender, and tumor pathologic stage was investigated in the TCGA cohort using Pearson correlation analysis and analysis of variance (ANOVA), or their non-parametric counterpart (Spearman correlation and Kruskal–Wallis test). The association between CD3D, CD4, and CD8A expression, ratios of CD3D/CD4 and CD3D/CD8A, and OS was analyzed with univariate and multivariate (adjusted for age, gender, and pathologic stage) Cox proportional hazard model. For independent prognostic markers identified by multivariate analysis, Kaplan–Meier survival curves were drawn for the low vs high (cutoff by median and optimal threshold suggested by Cox regression tree analysis) group comparisons; and median survivals and P-values by log-rank test were also calculated. Subgroup survival analyses stratified by molecular subtypes based on transcriptome were conducted. The identified prognostic factors were further subjected to external validation and pooled analysis using the validation datasets (original gene expression values by microarray were log2 transformed before analysis). A P<0.05 was considered statistically significant. The R version 3.4.3 was used for all statistical analyses.

**Results**

**Distribution of investigated markers in TCGA cohort**

The distribution of CD3D and CD8A mRNA expression was similar with each other, and significantly lower than CD4 expression across the 407 TCGA MIBC patients (mean difference (MD) =-2.21, 95% CI =[-2.30, -2.12], between CD3D and CD4; MD =-1.82, 95% CI =[-1.91, -1.72], between CD8A and CD4; Figure 1A). The CD3D/
CD8A ratio was significantly larger than CD3D/CD4 ratio (MD =0.31, 95% CI = [0.28, 0.33]; Figure 1B).

Correlation between investigated markers and clinical parameters in TCGA cohort

According to Pearson correlation analysis, the age was not correlated with any of the investigated markers (CD3D: \( r = -0.04, P = 0.43 \); CD4: \( r = -0.005, P = 0.9171 \); CD8A: \( r = 0.005, P = 0.92 \); CD3D/CD4: \( r = -0.07, P = 0.17 \); CD3D/CD8A: \( r = -0.07, P = 0.15 \)). A significant gender difference was detected only for CD4 (ANOVA, \( P = 0.037 \)) and CD3D/CD8A (Kruskal–Wallis test, \( P = 0.0001 \)). No significant association was found between pathologic stage (II, III, and IV) and any markers (CD3D, ANOVA, \( P = 0.40 \); CD4, ANOVA, \( P = 0.67 \); CD8A, ANOVA, \( P = 0.054 \); CD3D/CD4, Kruskal–Wallis test, \( P = 0.18 \); CD3D/CD8, Kruskal–Wallis test, \( P = 0.70 \)).

Prognostic significance of investigated markers in TCGA cohort

According to the findings of univariate Cox regression analysis, the CD3D and CD3D/CD4 ratio were significantly associated with OS (Table 1). Multivariate Cox regression adjusted for age, gender and pathologic stage (II vs III–IV) indicated that both the CD3D and CD3D/CD4 ratio were independent prognostic factors (CD3D, beta = -0.11, \( P = 0.032 \); CD3D/CD4, beta = -0.87, \( P = 0.025 \); Table 1). The CD3D/CD4 ratio had a better prognosis predictability than CD3 expression as measured by the concordance index in both univariate (0.586 vs 0.57) and multivariate models (0.657 vs 0.655).

We then divided the subjects into low vs high groups according to the expression of CD3D and CD3D/CD4 ratio, using median and the optimal cutoffs identified by Cox regression tree analysis. The CD3D high group showed non-significantly better OS than the CD3D low group when using median cutoff (median OS: 1,348 days vs 819 days, log-rank test \( P = 0.072 \); Figure 2A), but the difference became significant taking the optimal cutoff by 73.4% (median OS: 2,828 days vs 819 days, log-rank test \( P = 0.00056 \); Figure 2B). In contrast, the CD3D/CD4 high group showed significantly better OS than the CD3D/CD4 low group with both the median (median OS: 1,804 days vs 690 days, log-rank test \( P = 0.00055 \); Figure 2C) and the optimal cutoff by 59.6% (median OS: 2,828 days vs 680 days, log-rank test \( P = 4E-06 \); Figure 2D).

Subgroup analyses by molecular subtypes based on transcriptome for CD3D and CD3D/CD4 in TCGA cohort

It had been shown that some investigated TILs markers (precisely CD8A, CXCL9, and CXCL10) were unequally distributed among previously defined different TCGA molecular subtypes based on transcriptome.\(^{10,20}\) Therefore, the subgroup-specific prognostic significance of CD3D and CD3D/CD4 was further investigated by univariate and multivariate Cox regression analysis within these subtypes. As shown in Table 2, the CD3 expression and CD3D/CD4 ratio showed no significance in luminal-papillary, luminal and neuronal subtypes, and displayed limited prognostic predictability in luminal-infiltrated subtype. In contrast, the two markers demonstrated very strong
and CD3D/CD4 was stronger among subjects possibly classified as basal-squamous tumors in these validation cohorts. We considered subjects whose average expression level of the 10 key basal-squamous markers reported in Robertson et al.\(^9\) was within the top third as having basal-squamous tumors. However, no significant association with OS was observed (Table 3).

### Pooled analysis for CD3D/CD4 ratio across TCGA and validation cohorts

Random-effects meta-analysis was performed to combine the individual analysis results across TCGA and validation cohorts for CD3D/CD4 ratio. As shown in Figure 4, the pooled analysis indicated that considering all observations of included datasets, there was a significant association between CD3D/CD4 ratio and OS (beta = -1.31, 95% CI = [-2.23, -0.38]).

### Discussion

BC is a common and aggressive cancer that represents significant disease burden. Treatment options regarding MIBC are still limited. Although emerging active immunotherapy (anti-PD1/PD-L1) has shown promising efficacy in BC, only a subset of patients respond well. Besides, it is still largely unknown the way to predict/identify these responders and the reason why some MIBC molecular subtypes have poor TILs. On the other hand, TIL therapy, a passive immunotherapy using infusions of ex vivo expanded self-TILs to trigger anti-tumor immunity, has achieved great success over the past 10–15 years\(^9,21\) and actively studied in clinical trials involving several cancer types including melanoma, NSCLC and cervical cancer (NCT02360579; NCT03419559 and NCT03108495). Regarding TMB, which provide a preferential niche for immunotherapy, BC ranked third just following melanoma and NSCLC. In view of this, TIL biomarkers with prognostic significance may also be used to help selecting patients who may respond to immunotherapy and improve the clinical management of MIBC.

The cancer patients’ outcome is determined by the interaction between tumor and surrounding microenvironment. Previous studies have revealed that the microenvironment plays an important role in BC carcinogenesis and progression. Of note, it has been reported that TILs and their subpopulation structure are associated with prognosis in BC patients. Like MIBC tumors which are heterogeneous in terms of molecular profiles, the TILs are also a heterogeneous group of immune cells. The majority of

---

### Table 1 Prognostic significance of TIL markers in MIBC (TCGA, n=406)

<table>
<thead>
<tr>
<th>Markers (^a)</th>
<th>Univariate</th>
<th>Multivariate(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beta</td>
<td>P</td>
</tr>
<tr>
<td>CD3D</td>
<td>-0.11</td>
<td>0.025(^a)</td>
</tr>
<tr>
<td>CD4</td>
<td>-0.04</td>
<td>0.57</td>
</tr>
<tr>
<td>CD8A</td>
<td>-0.09</td>
<td>0.065</td>
</tr>
<tr>
<td>CD3D/CD4</td>
<td>-0.99</td>
<td>0.013(^a)</td>
</tr>
<tr>
<td>CD3D/CD8A</td>
<td>-0.19</td>
<td>0.46</td>
</tr>
</tbody>
</table>

Notes: \(^a\)Gene expression measured as log2 RSEM, or derived ratio, as continuous variable. \(^b\)Multivariate Cox regression adjusted by age, gender and tumor stage.

Statistically significant.

Abbreviations: MIBC, muscle-invasive bladder cancer; TIL, tumor-infiltrating lymphocytes; TCGA, the Cancer Genome Atlas.
TILs belong to the T cell population, which are mostly CD3+. Cytotoxic T cells (CTLs), with CD8+ as a specific marker, are the main mediator of the host’s anti-cancer immunity. The activity of CD8+ CTLs depends on a complicated interaction network involving other T cell subtypes including T-helper cells (CD4+) and regulatory T cells (Treg), and other immune cell populations such as tumor-associated macrophages (TAMs). One of the first investigations on the role of TILs in MIBC by Sharma and colleagues revealed that higher infiltration of CD8+ CTLs was associated with better survival. Sjodahl et al found that presence of CD3+ T cell infiltration predicted good prognosis in MIBC patients, which involved

---

**Figure 2** Kaplan–Meier curves showing difference in OS in MIBC patients between (A) low and high CD3D groups by median cutoff; (B) low and high CD3D groups by optimal cutoff; (C) low and high CD3D/CD4 groups by median cutoff; (D) low and high CD3D/CD4 groups by optimal cutoff.

**Abbreviation:** MIBC, muscle-invasive bladder cancer.
modulation by CD68+ TAMs. Horn et al reported ratio between FOXP3 and CD3D or CD8 positive cells were significantly associated with BC survival. However, others found lack of evidence for a prognostic value of TILs in BC through a global pan-cancer analysis with TCGA data. We supposed that this inconsistency could be a result of ignoring the heterogeneous immune-infiltration status across different MIBC molecular subtypes. According to Mariathasan and colleagues, the immune-infiltration markers were enriched in luminal-infiltrated and basal-squamous subtypes compared with other subtypes. This can be partially explained by the PPARγ/RXRα pathway, which is mainly activated in luminal-papillary and luminal tumors, and confers immune evasion through impairing CD8+ T-cell infiltration. Accordingly, it has been shown that tumors belong to the luminal-infiltrated or basal-squamous subtypes respond better to immunotherapy than other subtypes.

Despite the aforementioned research efforts concerning the role of TILs in MIBC, the prognostic predictability of mRNA expression of major TIL markers (CD3D, for overall T cells; CD4, for T-helpers; and CD8A, for CTLs) and their ratios have not been reported. The immune infiltration status was usually measured by immunohistochemistry or flow cytometry counting of infiltrated immune cells. Without doubt, both of them were precise but time-consuming and technically complicated. Biomarkers based on gene expression are possible alternatives. Moreover, mRNA detection is more feasible to take advantage of liquid-based urine cytology, which represents the “gold standard” for surveillance of urothelial carcinoma diagnosis and prognosis.

In the TCGA investigation cohort, we observed that the CD4 expression was significantly higher than CD3D expression. The CD3D/CD4 ratio was less than 1 in most MIBC subjects, and CD3D/CD8A less than 1 in more than half of the subjects. It seems that in many of the subjects the TILs had a large number of and/or were actively proliferating CD3+ CD4+ (macrophages, monocytes, and dendritic cells) and CD3+CD8+ cells (NK cells). We thus suppose that the CD3D/CD4 ratio can be considered a rough measurement of interaction between T cells and CD3+CD4+ cells which could be immune-suppressing cells, which is in accordance with Sjodhal et al’s findings. In line with previous reports, we also found CD8A was an independent prognostic factor but only in basal-squamous subtypes. Interestingly, the overall T cell marker CD3D and the ratio of CD3D/CD4 were independent favorable prognostic factors in the TCGA cohort; furthermore, a stable association was observed in several validation cohorts for CD3D/CD4, strengthening the reliability of this predictor. Strikingly, both CD3D and the ratio of CD3D/CD4 exhibited largely stronger predictability in TCGA patients with basal-squamous tumor; although the validation in other cohorts failed to identify a stable significant association which could be due to lack of consensus classification among different cohorts and very limited sample size. This suggests that the interaction between immune-effecting CD3+ TILs and immune-suppressing CD3+CD4+ cells may play a particularly important role in the anti-tumor immunity in basal-squamous tumors. Following this clue, we speculate that basal-squamous MIBC patients with higher expression ratio of CD3D/CD4 (stronger effecting TILs and weaker suppressing cells) could be potential responders to cancer immunotherapy, which should be further investigated.

In conclusion, we have found that high CD3D/CD4 ratio predicts better survival in MIBC, and its association with OS may be stronger in basal-squamous tumors.
Table 3 | Results of validation analyses in three independent cohorts

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Overall subjects</th>
<th>Subjects with high basal-squamous markers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>CD3D&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>CIT</td>
<td>73</td>
<td>−0.024</td>
</tr>
<tr>
<td>GSE31684</td>
<td>78</td>
<td>−0.21</td>
</tr>
<tr>
<td>GSE5287</td>
<td>30</td>
<td>−0.26</td>
</tr>
</tbody>
</table>

Notes: *Gene expression measured as log2 RSEM, or derived ratio, as continuous variable. **Statistically significant.

Figure 3 Kaplan–Meier curves showing difference in OS in basal-squamous MIBC patients between (A) low and high CD3D groups by median cutoff; (B) low and high CD3D groups by optimal cutoff; (C) low and high CD3D/CD4 groups by median cutoff; (D) low and high CD3D/CD4 groups by optimal cutoff.

Abbreviation: MIBC, muscle-invasive bladder cancer.
A hypothesis is made that CD3D/CD4 may also predict response to immunotherapy in basal-squamous tumors, which needs further investigation.

Acknowledgment
This work has been supported by the Fundamental Research Funds for the Central Universities (2662018PY023).

Disclosure
The authors report no conflicts of interest in this work.

References
### Table S1 Basic characteristics of included datasets

<table>
<thead>
<tr>
<th></th>
<th>Investigation</th>
<th>Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TCGA (n=407)</td>
<td>CIT (n=73)</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>Percentage</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤65</td>
<td>160</td>
<td>39.3%</td>
</tr>
<tr>
<td>&gt;65</td>
<td>246</td>
<td>60.4%</td>
</tr>
<tr>
<td>Unknown</td>
<td>1</td>
<td>0.2%</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>300</td>
<td>73.7%</td>
</tr>
<tr>
<td>Female</td>
<td>107</td>
<td>26.3%</td>
</tr>
<tr>
<td><strong>Tumor stage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>132</td>
<td>32.4%</td>
</tr>
<tr>
<td>III/IV</td>
<td>273</td>
<td>67.1%</td>
</tr>
<tr>
<td>Unknown</td>
<td>2</td>
<td>0.5%</td>
</tr>
<tr>
<td><strong>Molecular subtype</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Luminal Papillary</td>
<td>136</td>
<td>33.4%</td>
</tr>
<tr>
<td>Luminal Infiltrated</td>
<td>76</td>
<td>18.7%</td>
</tr>
<tr>
<td>Luminal</td>
<td>43</td>
<td>10.6%</td>
</tr>
<tr>
<td>Basal-Squamous</td>
<td>133</td>
<td>32.7%</td>
</tr>
<tr>
<td>Neuronal</td>
<td>19</td>
<td>4.7%</td>
</tr>
<tr>
<td><strong>Adjuvant chemotherapy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>176</td>
<td>43.2%</td>
</tr>
<tr>
<td>No</td>
<td>182</td>
<td>44.7%</td>
</tr>
<tr>
<td>Unknown</td>
<td>149</td>
<td>36.6%</td>
</tr>
<tr>
<td><strong>Adjuvant radiotherapy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>10</td>
<td>2.5%</td>
</tr>
<tr>
<td>No</td>
<td>263</td>
<td>64.6%</td>
</tr>
<tr>
<td>Unknown</td>
<td>134</td>
<td>32.9%</td>
</tr>
</tbody>
</table>

**Abbreviation:** NA, not available.